

**Universidade de São Paulo  
Escola Superior de Agricultura “Luiz de Queiroz”**

**Abordagem bayesiana de modelos de regressão com platô de  
resposta**

**Rick Anderson Freire Mangueira**

Tese apresentada para obtenção do título de Doutor em Ciências. Área de concentração: Estatística e Experimentação Agronômica

**Piracicaba  
2019**

Rick Anderson Freire Mangueira  
Bacharel em Estatística

**Abordagem bayesiana de modelos de regressão com platô de resposta**  
versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientadora:  
Prof <sup>a</sup> Dr<sup>a</sup> **TACIANA VILLELA SAVIAN**

Tese apresentada para obtenção do título de Doutor em Ciências. Área de concentração: Estatística e Experimentação Agronômica

**Piracicaba**  
**2019**

**Dados Internacionais de Catalogação na Publicação  
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP**

Mangueira, Rick Anderson Freire  
Abordagem bayesiana de modelos de regressão com platô de resposta /  
Rick Anderson Freire Mangueira. – versão revisada de acordo com a resolução  
CoPGr 6018 de 2011 – Piracicaba, 2019.  
90 p.

Tese (Doutorado) – USP / Escola Superior de Agricultura “Luiz de  
Queiroz”.

. 1. Exigência nutricional 2. Platô de resposta 3. Regressão segmentada I.  
Título.

## AGRADECIMENTOS

Primeiramente agradeço a Deus, que está acima de tudo e todos, guiou meu caminho até o fim deste momento de minha jornada.

Agradeço à minha família, meus pais Sérgio e Mauricéia, irmãos Ryan e Raylla, por sempre estarem preocupados comigo enquanto estive na missão que foi esse doutorado.

À Gabriela, minha companheira durante todo período de doutorado.

À minha orientadora Taciana, que foi, na verdade, uma colega de trabalho, que sempre esteve disponível para conversar, tirar minhas dúvidas e contribuir com o trabalho.

Aos Amigos, docentes e funcionários do PPGEEA da ESALQ/USP.

Aos amigos piracicabanos, Adilson, Armando, Chico, Douglas, Horácio, Irineu, Odair, Regina, Renato, Roberto e o Seung, por permitirem momentos de alegria e descontração.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, que deu o suporte financeiro durante toda pesquisa.

## SUMÁRIO

RESUMO . . . . .	6
ABSTRACT . . . . .	7
LISTA DE FIGURAS . . . . .	8
LISTA DE TABELAS . . . . .	10
1 INTRODUÇÃO . . . . .	13
2 REVISÃO BIBLIOGRÁFICA . . . . .	15
2.1 Exigências nutricionais . . . . .	15
2.2 Modelos de regressão segmentada . . . . .	15
2.3 Modelos de regressão segmentada com platô de resposta . . . . .	19
2.4 Abordagem bayesiana . . . . .	23
2.4.1 Distribuições a priori e a posteriori . . . . .	24
2.4.2 Método de Monte Carlo via Cadeia de Markov (MCMC) . . . . .	25
2.4.3 Amostrador de Gibbs . . . . .	26
2.4.4 Metropolis-Hastings . . . . .	27
2.4.5 Convergência das cadeias . . . . .	27
2.4.6 Intervalo de Credibilidade - HPD . . . . .	28
3 MATERIAL E MÉTODOS . . . . .	31
3.1 Material . . . . .	31
3.1.1 Simulação . . . . .	31
3.1.2 Dados experimentais . . . . .	33
3.2 Método . . . . .	34
3.2.1 Modelos empíricos . . . . .	34
3.2.2 Metodologia bayesiana . . . . .	36
3.2.3 Diagnóstico e critério de comparação dos ajustes . . . . .	41
4 RESULTADOS E DISCUSSÃO . . . . .	43
4.1 Estudo de simulação para o modelo de regressão quadrático com platô de resposta (MQP) . . . . .	43
4.2 Estudo de simulação para o modelo não linear com platô de resposta (MNL) . . .	47
4.3 Análise dos dados de teor de Zinco (Zn) depositados na tíbia de frangos . . . . .	50

4.3.1 Ajuste do modelo de regressão quadrático com platô de resposta (MQP) e do modelo não linear com platô de resposta(MNL), considerando variância homogênea	52
4.3.2 Ajuste do modelo de regressão quadrático com platô de resposta (MQP) e do modelo não linear com platô de resposta(MNL), considerando heterogeneidade de variâncias . . . . .	55
4.3.3 Diagnósticos dos ajustes e critério de seleção do modelo . . . . .	58
5 CONCLUSÃO . . . . .	61
REFERÊNCIAS . . . . .	63
ANEXOS . . . . .	67

## RESUMO

### Abordagem bayesiana de modelos de regressão com platô de resposta

Estudos para estabelecer níveis de nutrientes adequados em dietas de animais são bastante comuns no meio científico. Tais estudos são realizados utilizando a média da variável resposta nos diferentes níveis da variável independente e, geralmente, por meio de abordagem frequentista. Sendo assim, muita informação é perdida ou até mesmo as médias influenciadas por valores extremos nos resultados das repetições. O objetivo do presente trabalho foi realizar um estudo de simulação de modelos de regressão com platô de resposta, considerando diversos cenários de diferentes tamanhos amostrais, além da homogeneidade e heterogeneidade de variância entre níveis do fator dosagem levando em consideração as repetições em cada dosagem, por meio de abordagem bayesiana. Também foi realizado ajustes dos modelos de regressão quadrático com platô de resposta e do não linear com platô de resposta utilizando a metodologia bayesiana para estimar os parâmetros dos modelos. Os dados simulados representavam o teor de Zinco (Zn) depositado na tíbia de frangos, que receberam as dosagens 0, 15, 30, 45, 60, 75, 90, 105 e 120 ppm de Zn na ração. Nos cenários 1, 2 e 3 foi atribuído variância constante, com tamanhos amostrais de  $n=16$ ,  $n=30$  e  $n=100$ , respectivamente. Nos cenários 4, 5 e 6, foram considerados também os tamanhos amostrais  $n=16$ ,  $n=30$  e  $n=100$ , respectivamente, mas com diferentes variâncias ao longo do fator dose. Os dados utilizados para o estudo sobre os modelos de regressão com platô de resposta são referentes ao teor de Zinco (Zn), em ppm, na tíbia de frangos e frangas de corte como variável resposta. A análise de simulação mostrou que é adequada a utilização dos modelos com platô de resposta considerando a repetição das amostras em cada dose para estimar o nível adequado de um nutriente na ração animal. No ajuste dos modelos aos dados reais, o modelo de regressão quadrático com platô de resposta considerando a heterogeneidade de variâncias foi o que apresentou melhor ajuste, em que obteve-se uma estimativa de  $\hat{x}_0 = 73,26$  ppm para o nível de Zn aplicado na ração para uma deposição de Zn na tíbia dos frangos de  $\hat{P} = 290,12$  ppm.

Palavras-chave: Exigência nutricional, Platô de resposta, Regressão segmentada.

## ABSTRACT

### Bayesian approach of regression models with response plateau

Studies to establish adequate levels of nutrients in animal diets are quite common in the scientific world. Such studies are performed using the mean of the response variable in the different levels of the independent variable and, generally, through a frequentist approach. Thus, much information is lost or even the means influenced by extreme values in the results of the repetitions. The objective of the present work was to perform a simulation study of regression models with response plateau, where as different scenarios of different sample sizes, besides the homogeneity and heterogeneity of variance between levels of the dosage factor taking the repetitions in each dosage into account, through Bayesian approach. Fits were also made to the quadratic regression models with response plateau and nonlinear regression with response plateau using Bayesian methodology to estimate the parameters of the models. The simulated data represented the zinc content (Zn) deposited in the tibia of broilers, which received the dosages 0, 15, 30, 45, 60, 75, 90, 105 and 120 ppm of Zn in the feed. In the scenarios 1, 2 and 3 were attributed constant variance with sample sizes of  $n = 16$ ,  $n = 30$  and  $n = 100$ , respectively. In the scenarios 4, 5 and 6 the sample sizes were, respectively,  $n = 16$ ,  $n = 30$  and  $n = 100$ , but with different variances along the dose factor. The data used for the study on the regression models with response plateau refer to the Zn content, in ppm, in the tibia of male and female broilers as response variable. The simulation analysis showed that it is appropriate to use the models with response plateau considering the repetition of the samples at each dose to estimate the adequate level of a nutrient in the animal feed. In the fit of the models to the real data, the quadratic regression model with response plateau considering the heterogeneity of variances was the one that presented the best fit, which an estimate of  $\hat{x}_0 = 73.26$  ppm for the level of Zn applied to the feed for a Zn deposition in the chicken tibia of  $\hat{P} = 290.12$  ppm.

Keywords: Nutritional requirement, response plateau, Segmented regression.



## LISTA DE FIGURAS

Figura 1 - Valores paramétricos utilizados no estudo de simulação por meio do modelo MQP . . . . .	32
Figura 2 - Valores paramétricos utilizados no estudo de simulação por meio do modelo MNL . . . . .	32
Figura 3 - Boxplot das estimativas e vieses dos parâmetros $\alpha$ , $\beta$ e $\gamma$ do modelo de regressão quadrático com platô de resposta em cada cenário simulado . . .	46
Figura 4 - Boxplot das estimativas e vieses dos parâmetros $x_0$ e P do modelo de regressão quadrático com platô de resposta em cada cenário simulado . . .	47
Figura 5 - Estimativas dos parâmetros do modelo não linear com platô de resposta em cada cenário simulado . . . . .	49
Figura 6 - Boxplot para teor de Zn na tíbia de frangos de corte. . . . .	50
Figura 7 - Boxplot para teor de Zn na tíbia de frangos de corte, sem os outliers. . . .	51
Figura 8 - Ajuste dos modelos MQP e MNL, aos dados do teor de Zinco na tíbia de frangos, no caso em que foi considerado homogeneidade de variância, e intervalos de credibilidade (linhas verticais) para o ponto de troca . . . . .	55
Figura 9 - Ajuste dos modelos MQP e MNL, aos dados do teor de Zinco na tíbia de frangos, no caso em que foi considerado heterogeneidade de variâncias, e intervalos de credibilidade (linhas verticais) para o ponto de troca. . . . .	58
Figura 10 - Resíduos padronizados x valores preditos do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta para os casos em que considerou-se homogeneidade de variância. . . . .	59
Figura 11 - Resíduos padronizados x valores preditos do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta para os casos em que considerou-se heterogeneidade de variâncias. . . . .	59
Figura 12 - Histograma das estimativas dos parâmetros do modelo de regressão quadrático com platô de resposta para os cenários 1, 2 e 3 simulados. . . . .	76
Figura 13 - Estimativas dos parâmetros do modelo de regressão quadrático com platô de resposta, em função do tamanho amostral na simulação. . . . .	77
Figura 14 - Histograma das estimativas dos parâmetros do modelo de regressão quadrático com platô de resposta para os cenários 4, 5 e 6 simulados. . . . .	78

Figura 15 - Estimativas dos parâmetros do modelo de regressão quadrático com platô de resposta, em função do tamanho amostral na simulação. . . . .	79
Figura 16 - Histograma das estimativas dos parâmetros do modelo não linear com platô de resposta para os cenários 1, 2 e 3 simulados. . . . .	80
Figura 17 - Estimativas dos parâmetros do modelo não linear com platô de resposta, em função do tamanho amostral na simulação. . . . .	81
Figura 18 - Histograma das estimativas dos parâmetros do modelo não linear com platô de resposta para os cenários 4, 5 e 6 simulados. . . . .	82
Figura 19 - Estimativas dos parâmetros do modelo não linear com platô de resposta, em função do tamanho amostral na simulação. . . . .	83
Figura 20 - Cadeias dos processos numéricos para obtenção das distribuições a posteriori para parâmetros do MQP com homogeneidade de variância. . . . .	84
Figura 21 - Densidades a posteriori dos parâmetros do MQP com homogeneidade de variância. . . . .	85
Figura 22 - ACF dos parâmetros do MQP com homogeneidade de variância. . . . .	85
Figura 23 - Cadeias dos processos numéricos para obtenção das distribuições a posteriori para parâmetros do MNL com homogeneidade de variância. . . . .	86
Figura 24 - Densidades a posteriori dos parâmetros do MNL com homogeneidade de variância. . . . .	86
Figura 25 - ACF dos parâmetros do MNL com homogeneidade de variância. . . . .	87
Figura 26 - Cadeias dos processos numéricos para obtenção das distribuições a posteriori para parâmetros do MQP com heterogeneidade de variância. . . . .	87
Figura 27 - Densidades a posteriori dos parâmetros do MQP com heterogeneidade de variância. . . . .	88
Figura 28 - ACF dos parâmetros do MQP com heterogeneidade de variância. . . . .	88
Figura 29 - Cadeias dos processos numéricos para obtenção das distribuições a posteriori para parâmetros do MNL com heterogeneidade de variância. . . . .	89
Figura 30 - Densidades a posteriori dos parâmetros do MNL com heterogeneidade de variância. . . . .	89
Figura 31 - ACF dos parâmetros do MNL com heterogeneidade de variância. . . . .	90

## LISTA DE TABELAS

Tabela 1 - Médias das estimativas da média a posteriori dos parâmetros do modelo de regressão quadrático com platô de resposta e vieses médios, para os cenários 1, 2 e 3. . . . .	43
Tabela 2 - Médias das estimativas da média a posteriori dos parâmetros do modelo de regressão quadrático com platô de resposta e vieses médios, para os cenários 4, 5 e 6. . . . .	44
Tabela 3 - Médias das estimativas da média a posteriori dos parâmetros de variância do modelo de regressão quadrático com platô de resposta e vieses médios, nos cenários 4, 5 e 6. . . . .	44
Tabela 4 - Médias das estimativas da média a posteriori dos parâmetros do modelo não linear com platô de resposta e vieses médios, para os cenários 1, 2 e 3. . . . .	48
Tabela 5 - Médias das estimativas da média a posteriori dos parâmetros do modelo não linear com platô de resposta e vieses médios, para os cenários 4, 5 e 6. . . . .	48
Tabela 6 - Médias das estimativas da média a posteriori dos parâmetros de variância do modelo não linear com platô de resposta e vieses médios, nos cenários 4, 5 e 6. . . . .	49
Tabela 7 - Médias e desvios padrões do teor de Zn na tíbia dos frangos em cada dose de Zn. . . . .	52
Tabela 8 - Teste de convergência de cadeia de Geweke, modelo de regressão quadrático com platô de resposta e modelo não linear com platô de resposta, considerando homogeneidade de variância. . . . .	53
Tabela 9 - Médias a posteriori, desvios padrões (DP) a posteriori e intervalo de credibilidade HPD e erro de Monte Carlo (EMC) dos parâmetros do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta, considerando homogeneidade de variância. . . . .	54
Tabela 10 - Teste de convergência de cadeia de Geweke, modelo de regressão quadrático com platô de resposta e modelo não linear com platô de resposta, considerando heterogeneidade de variâncias. . . . .	56

Tabela 11 -Médias a posteriori, desvios padrões (DP) a posteriori, intervalo de credibilidade HPD e erro de Monte Carlo (EMC), dos parâmetros do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta, considerando uma variância para cada nível de dose. . .	57
Tabela 12 -Critério de informação deviance (DIC) dos modelos MQP e MNL. . . . .	60
Tabela 13 -Médias das estimativas dos parâmetros de precisão do modelo de regressão quadrático com platô de resposta e vieses médios nos ajustes dos cenários 4, 5 e 6. . . . .	74
Tabela 14 -Médias das estimativas dos parâmetros de precisão do modelo não linear com platô de resposta e vieses médios dos ajustes para os cenários 4, 5 e 6.	75



## 1 INTRODUÇÃO

As relações entre duas variáveis, sendo uma variável resposta e outra variável independente, podem apresentar diferentes estruturas em modelos estatísticos. Os modelos de regressão segmentada apresentam, na relação dessas variáveis, uma quebra ou desvio de direção significativo no decorrer da relação, ou seja, primeiramente há uma estrutura acompanhando a relação das variáveis e, a partir de certo ponto, há uma mudança considerável na direção da mesma.

Uma das classes de modelos de regressão segmentada utilizados em pesquisas de diversas áreas, são os modelos de regressão segmentada com platô de resposta. Esses modelos apresentam a característica de terem uma estrutura no modelo em que a relação das variáveis é uma reta constante paralela ao eixo das abscissas (platô de resposta), o que representa a estabilização da variável resposta a partir de um ponto da variável independente (ponto de troca).

Os modelos de regressão segmentada com platô de resposta são importantes para se obter um ponto em que haja a estabilização do crescimento (ou decrescimento) da variável resposta. Essa técnica é, geralmente, utilizada para serem obtidos tamanhos ótimos de parcelas em experimentos e, frequentemente, em estudos para identificar exigências nutricionais adequadas em rações para diversas espécies animais.

Na abordagem dos modelos na área de exigência nutricional, Coelho et al. (1987) destacam que os estudos que visam avaliar exigências nutricionais de animais consideram que a adição de um nutriente limitante em rações deficientes nesse nutriente, porém, adequada nos demais, resultará em crescimento do animal até que sua exigência seja atendida.

Identificar o ponto de estabilização do crescimento da variável resposta é de suma importância, pois pode trazer benefícios econômicos e melhor adequação do estudo realizado. Rostagno et al. (2007) afirmam a importância de manter um nível adequado de nutrientes nas rações para que não haja desperdícios econômicos e excesso nutricional nas dietas, pois o excesso de nutrientes nas dietas são eliminados nas fezes e urina, culminando na contaminação de solos e reservas de águas.

A estimação dos parâmetros dos modelos segmentados pode ser realizada por diversos métodos clássicos ou bayesianos. A utilização da abordagem bayesiana pode ser útil

para realizar inferências, obter resumos estatísticos do parâmetro de interesse e intervalos com uma dada probabilidade de ocorrência do verdadeiro valor do parâmetro.

Sendo assim, o objetivo do trabalho foi realizar um estudo dos modelos de regressão segmentada com platô de resposta obtendo estimativas dos parâmetros por abordagem bayesiana, realizando um estudo de simulação com cenários estruturados com diferentes tamanhos amostrais e diferentes informações sobre variância. Além disso, foram realizados ajustes de modelos com platô de resposta para dados reais do teor de Zinco depositados nas tíbias de frangos, também por meio de abordagem bayesiana.

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 Exigências nutricionais

Ter conhecimento dos níveis adequados de nutrientes em rações animais é de fundamental importância para economia e desenvolvimento do manejo das criações de animais, pois, segundo Rostagno et al. (2007) podem-se evitar os excessos de nutrientes nas dietas que são eliminados nas fezes e urina, resultando em contaminação de solos e de reservas de água. Além disso, conforme os mesmos autores, há um avanço a passos largos no tocante de exigências nutricionais, devido a competência e qualidade de pesquisas nesta área. Conforme Oliveira et al. (1999) o maior investimento na criação de aves é na alimentação, que corresponde a cerca de 75% do custo de produção final.

Ainda assim, segundo Gomes et al. (2009), é comum na suplementação em rações animais ser feita superior às exigidas, buscando um melhor desenvolvimento das aves. Isso ocorre devido a falta de conhecimento sobre o real valor de um nutriente na ração.

Tratando-se de exigências nutricionais para peixes, Portz, Dias e Cyrino (2000) afirmam que tal conhecimento é fundamental para a produtividade e economia dos sistemas de produções, além do aproveitamento eficiente dos nutrientes.

Lerman e Bie (1975) discutem duas grandes dificuldades para se alcançar dietas que maximizam os lucros agrícolas. A complexidade de predizer confiavelmente a resposta dos animais em uma dieta com qualquer composição e a variação do teor de nutrientes encontrada nos diferentes lotes de ração.

Com relação ao zinco (Zn), que é o mineral em estudo no presente trabalho, sua concentração está diretamente associada ao crescimento e ao desenvolvimento do tecido ósseo, além da proteção de membranas, efeito antioxidante que protege os grupos sulfidrilas nas membranas, metabolismo de prostaglandinas e metabolismo de lipídeos, assim é importante na formação esquelética e sua deficiência pode trazer anomalias de pernas e dedos (GOMES et al., 2008, 2009). Além disso, Bertechini, Hossain e Nobre (1992) afirmam que o melhor ponto de avaliação de exigência de Zn é encontrado na tíbia dos animais.

### 2.2 Modelos de regressão segmentada

Na literatura, é comum encontrar diferentes nomes para os modelos de regressão segmentada, como por exemplo *breakpoint*, *piecewise*, modelo *broken-line*, *change point*



e modelo de duas fases. Tal modelo é caracterizado por apresentar uma mudança estrutural ao longo da evolução da variável independente, ou seja, inicialmente, o modelo tem uma estrutura de direção, que pode ser crescente ou decrescente, e a partir de um ponto, denominado ponto de troca, há uma mudança nessa estrutura. Para Kaps e Lamberson (2004), o modelo de regressão segmentada é outra maneira de descrever uma relação curvilínea de uma variável dependente e outra independente.

Conforme apresentado por Brochi e Diniz (2001), uma forma geral do modelo de regressão segmentada pode ser definida como:

$$f(x; \boldsymbol{\theta}, \boldsymbol{\lambda}) = \begin{cases} f_1(x; \boldsymbol{\theta}_1), & x \leq \lambda_1; \\ f_2(x; \boldsymbol{\theta}_2), & \lambda_1 < x \leq \lambda_2; \\ \vdots & \vdots \\ f_D(x; \boldsymbol{\theta}_D), & \lambda_D < x, \end{cases}$$

sendo  $\boldsymbol{\theta}$  um vetor de parâmetros, D o número de segmentos que definem a função, e  $\boldsymbol{\lambda}$  o vetor de valores que delimitam os segmentos.

Existem várias parametrizações para escrever o modelo de regressão segmentada. Cada parametrização considera a continuidade ou não do modelo no ponto de troca de estrutura, sendo assim, o modelo de regressão segmentada mais simples, é o modelo linear-linear, ou seja, a estrutura é dada por dois segmentos de reta, que apresentam continuidade no ponto de troca. Sua equação é dada por:

$$Y_i = \begin{cases} \alpha_1 + \beta_1 x_i + \varepsilon_i, & x_i \leq \lambda; \\ \alpha_2 + \beta_2 x_i + \varepsilon_i, & x_i > \lambda. \end{cases}$$

em que,  $Y_i$  é a variável resposta, com  $i = 1, \dots, n$ ,  $x_i$  é a variável independente,  $\alpha_1$  e  $\alpha_2$  são os interceptos no eixo y do primeiro e segundo segmentos de reta, respectivamente,  $\beta_1$  e  $\beta_2$  são os coeficientes angulares da primeira e segunda retas, respectivamente,  $\lambda$  é o valor de  $x_i$  que ocorre o ponto de troca de estrutura, ou seja, a junção dos segmentos, e  $\varepsilon_i$  é o erro.

Robbins (1986) apresenta dois modelos segmentados denominados de *broken-line* de uma inclinação e de duas inclinações, representados, respectivamente, por:

$$Y = L + U(R - X_{LR}) + \varepsilon$$

e

$$Y = L + U(R - X_{LR}) + V(X_{GR} - R) + \varepsilon,$$

sendo L e R a ordenada e a abscissa do ponto de troca na curva, respectivamente; U é a inclinação da reta para  $X < R$  e V é a inclinação da reta quando  $X > R$ ;  $X_{LR}$  e  $X_{GR}$  são os valores de X quando  $X < R$  e  $X > R$ , respectivamente. Por definição,  $(R - X_{LR})$  é zero quando  $X > R$  e  $(X_{GR} - R)$  é zero em  $X < R$ .

Outro modelo de regressão segmentada encontrado na literatura é o apresentado por Quandt (1960). Nesse modelo é estimada a posição em que ocorre o ponto de troca, e não exatamente o valor que representa o ponto de troca. Então, tem-se:

$$Y_i = \begin{cases} \alpha_1 + \beta_1 x_i + \varepsilon_i, & i = 1, \dots, t; \\ \alpha_2 + \beta_2 x_i + \varepsilon_i, & i = t + 1, \dots, T, \end{cases}$$

sendo  $Y_i$  variável resposta,  $x_i$  variável independente, T são números de pares independentes de  $(x_i, Y_i)$  e  $\varepsilon_i$  é o erro do modelo. O propósito, nesse caso, está em estimar a posição t.

Segundo Brochi e Diniz (2001), um modelo de regressão segmentada é contínuo quando, considerando  $x_0$  o ponto de troca dos segmentos, tem-se as seguintes propriedades:

- i.  $\exists f(x_0)$ ;
- ii.  $\exists \lim_{x \rightarrow x_0} f(x)$ ;
- iii.  $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ .

Os modelos de regressão segmentada são utilizados em diversas áreas do conhecimento. Shaban (1980) faz uma revisão de vários estudos teóricos e aplicados importantes, referentes a regressão segmentada e ponto de troca, dentre eles, Barten e Bronsard (1970) que tratam da estimação de parâmetros de uma regressão de dois estágios, aplicando a metodologia a dados do mercado da melancia e Goldfeld e Quandt (1973) que analisam algumas técnicas econométricas para estudar modelos de regressão com ponto de troca.

Conforme Holbert (1982), as variáveis em estudo podem ser de natureza

econômica como uma mudança de política tributária ou um novo programa governamental, da área da biologia em que a ênfase está no crescimento de um organismo no qual ocorre uma mudança no padrão de crescimento ou até na geologia na relação de produtos químicos do solo abaixo de um local de perfuração.

Silva (1995b) apresenta fundamentos da regressão segmentada para identificar a caracterização da estabilidade fenotípica e adaptabilidade de plantas cultivadas. O autor esclarece que o modelo é mais apropriado do que um modelo de regressão linear simples para identificar genótipos com características desejáveis. A metodologia demonstrada em Silva (1995b) é ilustrada em Silva (1995a), o autor conclui que a regressão segmentada permitiu caracterizar e identificar o comportamento de resposta do genótipo em diferentes ambientes.

Em um estudo também para verificação da adaptabilidade de genótipos em ambientes, Silva (1998) fez uma adaptação no modelo segmentado que considera um ponto de troca dos segmentos, tornando o modelo não linear e mais flexível com possibilidade de melhor aproximação da relação das variáveis em estudo.

Schmittiel et al. (2017) utilizam o modelo segmentado para estudar a relação do índice de massa corporal de pacientes obesos avaliados via um programa telefônico de bem estar ao longo do tempo.

Os autores Chen, Ban-Weiss e Sanders (2018) utilizaram a regressão segmentada linear-linear para estudar a relação do consumo de energia e a temperatura ambiente em residências. Tal estudo foi importante para se obter, por meio do ponto de troca de segmentos, a temperatura estacionária em que ocorre o mínimo da relação consumo x temperatura, e não há necessidade de aquecimento ou refrigeração do ambiente.

Os trabalhos citados anteriormente, referem-se principalmente a modelos de regressão segmentada do tipo linear-linear, ou seja, as duas estruturas, anterior e posterior ao ponto de troca, são descritas por uma reta.

A partir do conhecimento das diversas parametrizações e casos dos modelos de regressão segmentada, um caso especial de tais modelos são os modelos de regressão segmentada com platô de resposta, em que uma das estruturas do segmento, geralmente, após o ponto de troca, é uma constante paralela ao eixo das abscissas.

## 2.3 Modelos de regressão segmentada com platô de resposta

Os modelos de regressão segmentada com platô de resposta são amplamente utilizados para determinar exigências nutricionais adequadas em dietas animais, bem como na estimativa do tamanho ótimo de parcelas para estudos experimentais. Esses modelos são caracterizados por apresentarem como estrutura, após o ponto de troca, um segmento constante, ou seja, um segmento paralelo ao eixo das abscissas que se refere a uma estabilização da variável resposta.

Para Schabenberger e Pierce (2002), a proporcionalidade entre Y e X antes de atingir o ponto do platô, implica em uma relação linear que pode ser modelada por uma regressão linear. Logo que o aumento linear é combinado com o platô, o modelo correspondente é não linear. Pesti et al. (2009) afirmam que diferentes curvas que relacionam a resposta animal à composição da ração podem ser ajustadas aos dados de experimentos de exigências nutricionais.

Os modelos de regressão segmentada com platô de resposta apresentam em sua primeira estrutura diversas formas e parametrizações, e podendo ser representados por modelos lineares ou não lineares em relação aos parâmetros. Alguns desses modelos encontrados em Rezende (2002) e Rezende et al. (2007) são dados por:

a) Modelo linear com platô de resposta (MPL)

$$Y_i = \begin{cases} \alpha + \beta x_i + \varepsilon_i & \text{se } x_i \leq x_0; \\ p + \varepsilon_i & \text{se } x_i > x_0. \end{cases}$$

b) Modelo polinomial quadrático com platô de resposta (MQP)

$$Y_i = \begin{cases} \alpha + \beta x_i + \gamma x_i^2 + \varepsilon_i & \text{se } x_i \leq x_0; \\ p + \varepsilon_i & \text{se } x_i > x_0. \end{cases}$$

c) Modelo não-linear exponencial 1 com platô de resposta (MNLE1)

$$Y_i = \begin{cases} \alpha \exp(-\gamma(x_i - \beta)^2) + \varepsilon_i & \text{se } x_i \leq x_0; \\ p + \varepsilon_i & \text{se } x_i > x_0. \end{cases}$$

d) Modelo não-linear exponencial 2 com platô de resposta (MNLE2)

$$Y_i = \begin{cases} \alpha \exp(\beta x_i - \gamma x_i^2) + \varepsilon_i & \text{se } x_i \leq x_0; \\ p + \varepsilon_i & \text{se } x_i > x_0. \end{cases}$$

e) Modelo logarítmico com platô de resposta (MLOG)

$$Y_i = \begin{cases} \beta \log(x_i + \alpha) + \varepsilon_i & \text{se } x_i \leq x_0; \\ p + \varepsilon_i & \text{se } x_i > x_0, \end{cases}$$

em que:  $Y_i$  é a variável resposta,  $i = 1, \dots, n$ ;  $x_i$  seria o  $i$ -ésimo valor da variável explicativa;  $\alpha$ ,  $\beta$  e  $\gamma$  são parâmetros do modelo cuja interpretação depende do modelo utilizado;  $x_0$  é o valor da variável explicativa em que ocorre a mudança da estrutura do modelo;  $p$  é o platô de resposta, o máximo (ou mínimo) atingido pela variável resposta e  $\varepsilon_i$  é o erro associado ao ajuste do modelo.

Em cada um dos modelos apresentados, o parâmetro ponto de troca  $x_0$  é definido de uma forma diferente. Essa é uma importante informação, pois em um contexto de desenvolvimento animal, o nível ideal de um nutriente necessário para o máximo ganho de peso de uma espécie é dado pelo ponto de troca (PORTZ; DIAS; CYRINO, 2000).

O MPL é caracterizado por ter uma reta crescente ou decrescente no primeiro segmento, é o modelo mais simples dos modelos de regressão segmentada com platô de resposta. Já o MQP apresenta em sua primeira estrutura um segmento de parábola que em seu ponto de máximo ou mínimo encontra-se com o platô de resposta. Os demais modelos também apresentam uma curva no primeiro segmento que se intercepta com o platô de resposta.

Os modelos MPL e MQP foram utilizados por Peixoto, Faria e Moraes (2011) para estimar o tamanho ótimo de parcelas em experimentos de conservação do maracujazeiro *in vitro*. Os autores concluíram que o MQP seria mais adequado, pois foi o que se ajustou melhor em todos os tratamentos testados. Também para estimar tamanho ótimo de parcelas Moreira et al. (2016) ajustaram os MPL, MQP, além do método da curvatura máxima para experimentos para cultura do cafeeiro. Neste caso, o método da curvatura máxima mostrou-se mais adequado.

Guarconi et al. (2017) utilizaram o MPL, além do método da curvatura máxima modificado, para estimar o tamanho ótimo de parcela para experimentos com repolho, utilizando como variáveis resposta massa, diâmetro e compacidade do vegetal. Os autores concluíram que o MPL foi mais adequado na estimação das três variáveis resposta utilizadas, comparado com o método da curvatura máxima modificado.

Coldebella et al. (2005) utilizaram o MPL, um modelo de regressão linear simples e um modelo de regressão quadrática para determinar a estabilização da variabilidade

da energia metabolizável em pintos de corte. Os autores utilizaram critérios de comparação de modelos e concluíram que o melhor ajuste foi dado pelo MPL. Runho et al. (2001) também ajustaram os MPL e um modelo de regressão quadrático para obter a exigência nutricional de fósforo para frangos de corte machos e fêmeas, os autores comentam que o modelo quadrático superestimou os valores de exigência de fósforo e o MPL subestimou tal valor.

Portz, Dias e Cyrino (2000) utilizaram modelos segmentados de uma inclinação e de duas inclinações para verificar a exigência nutricional da dieta de peixes e comentam que esses modelos podem ser mais adequados que modelos não lineares, pois apresentam resultados mais coerentes.

Já Rezende et al. (2007) utilizaram os MQP, MNLE1 e MNLE2 para estimar a dose de zinco ideal nas rações de frangos de corte para se obter a deposição de zinco adequado na tíbia das aves. Os autores concluíram que o modelo MNLE1 seria recomendado por ser de melhor interpretação. Pesti et al. (2009) estudaram vários modelos para estimar o nível nutricional adequado para frangos de corte por meio da variável lisina. Dentre os modelos, o MPQ se destacou na região onde há a junção dos segmentos.

Para estimar as exigências de lisina em aves de corte, Siqueira et al. (2009) ajustaram o MPL com uma diferente parametrização, além dos modelos de regressão quadrático, exponencial e por fim o MQP que apresentou melhor desempenho para o objetivo de determinar o nível de aminoácidos.

Para determinar o nível total de lisina adequado para fêmeas de frangos de corte, Kim et al. (2018) ajustaram os modelos linear com platô e quadrático com platô. Os autores concluíram que 0,68% total de lisina na dieta, seria um teor adequado para as aves.

Todos os estudos citados anteriormente com modelos de regressão segmentada com platô de resposta tiveram as estimativas dos parâmetros obtidas por métodos clássicos como mínimos quadrados e máxima verossimilhança. A utilização de uma abordagem bayesiana não é comum nesses tipos de estudo, porém podem-se obter estimativas de parâmetros válidas por esse método.

A estimação dos parâmetros dos modelos de regressão segmentada pode ser realizada por diversos métodos clássicos, paramétricos e não-paramétricos ou por abordagem bayesiana.

Brochi e Diniz (2001) discutem alguns métodos de estimação para modelos

de regressão segmentada simples, tais como máxima verossimilhança, busca na grade e *bootstrap* considerando diferentes tamanhos amostrais e diferentes pontos de trocas, além de comentar sobre outros métodos de estimação como *piecewise* e não-paramétricos.

O método de máxima verossimilhança foi utilizado por Quandt (1958) para estimar os parâmetros de um modelo de regressão segmentada para caso de ocorrer apenas uma troca de estrutura. Já Quandt (1972) propõe uma nova abordagem na estimação dos parâmetros do modelo de regressão segmentada, em que se considera uma probabilidade desconhecida  $\lambda$  das observações pertencerem a um segmento e  $1 - \lambda$  de pertencerem ao outro segmento.

Lerman (1980) utiliza o método da busca na grade para estimar os parâmetros do modelo de regressão segmentada, ilustrando o método com três exemplos.

Um nova abordagem para estimação foi proposto em Muggeo (2003) para modelos segmentados com ponto de troca desconhecidos. Com essa abordagem, o autor soluciona problemas como de não linearidade e não diferenciabilidade.

A abordagem bayesiana também pode ser utilizada para estimar parâmetros de modelos de regressão segmentados. Bacon e Watts (1971) ilustram a estimação do ponto de troca de estrutura de duas fases por meio de uma abordagem bayesiana, assumindo *prioris* não informativas para os parâmetros.

Ferreira (1975) utilizou três distribuições a priori uniformes para o parâmetro ponto de troca do modelo segmentado com dois regimes. Além disso, o autor propõe intervalos de credibilidade baseados na distribuição t-Student e qui-quadrado para realização de testes de hipóteses para os parâmetros do modelo. Já ChinChoy e Broemeling (1980) utilizaram distribuição a priori gama-normal para os parâmetros do modelo segmentado obtendo distribuições a posteriori mistura de t-Student. Holbert (1982) também fez uso da abordagem bayesiana para modelos de regressão segmentada.

No trabalho de Carlin, Gelfand e Smith (1992), os autores obtiveram amostras das distribuições a posteriori marginais utilizando amostrador de Gibbs, solucionando assim as dificuldades encontradas na obtenção de resumos a posteriori. Diniz, Milan e Mazucheli (2003) também utilizaram o amostrador de Gibbs e de Metropolis-Hasting, além disso, os autores consideraram erros heterocedásticos e com distribuição t-Student multivariada.

Os métodos de estimação bayesiano, Julious (Julious (2001)), busca na grade e segmentado, foram avaliados por meio de simulação por Chen et al. (2011). Os autores

concluíram que apenas o método bayesiano foi adequado para modelos que apresentam ou não continuidade.

Um diferencial nos trabalhos citados anteriormente sobre estimação dos parâmetros dos modelos de regressão segmentada é que o ponto de troca estimado é a posição em que ocorre a troca de estrutura no modelo e não exatamente o valor da variável independente em que ocorre a mudança de estrutura ou inicia-se o platô de resposta. No presente trabalho o foco está na estimação do ponto de troca (valor da variável independente) em que ocorre a estabilização da variável resposta.

## 2.4 Abordagem bayesiana

Na metodologia bayesiana paramétrica combina-se a informação dos dados sobre os parâmetros, descrita pela função de verossimilhança, com o conhecimento prévio do pesquisador relativo aos parâmetros, descrito pela distribuição a priori. Segundo Paulino et al. (2018) a distribuição a priori expressa o grau de credibilidade que o pesquisador atribui ao parâmetro. A partir da combinação da informação a priori e dos dados, utilizando-se, o Teorema de Bayes, obtém-se a distribuição a posteriori.

O Teorema de Bayes

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_i P(B|A_i)P(A_i)}, \quad (1)$$

sendo  $A_i$  eventos que formam uma partição de um espaço amostral, com  $m$  eventos, e  $B$  é um evento qualquer.

A distribuição a posteriori é dada por:

$$\pi(\boldsymbol{\theta} | x) = \frac{L(\boldsymbol{\theta}|x) \times \pi(\boldsymbol{\theta})}{\pi(x)} = \frac{L(\boldsymbol{\theta}|x) \times \pi(\boldsymbol{\theta})}{\int L(\boldsymbol{\theta}|x) \times \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}}$$

em que:  $\boldsymbol{\theta}$  um vetor de parâmetros;  $\pi(\boldsymbol{\theta}|x)$  é a distribuição conjunta a posteriori do vetor de parâmetros dado os dados;  $L(\boldsymbol{\theta}|x)$  é a função de verossimilhança,  $\pi(\boldsymbol{\theta})$  é distribuição a priori conjunta para o vetor de parâmetros  $\boldsymbol{\theta}$  e  $\pi(x)$  é a distribuição marginal de  $x$ . Como  $\frac{1}{\pi(x)}$  não depende de  $\boldsymbol{\theta}$ , resulta em uma constante normalizadora de  $\pi(\boldsymbol{\theta} | x)$ , obtém-se, então, a expressão 2.



$$\pi(\boldsymbol{\theta} \mid x) \propto L(\boldsymbol{\theta} \mid x) \times \pi(\boldsymbol{\theta}). \quad (2)$$

### 2.4.1 Distribuições a priori e a posteriori

Para Gelman et al. (2014), a distribuição a priori representa o conhecimento sobre os possíveis valores para o parâmetro em estudo, deve incluir todos os valores plausíveis para o mesmo. A distribuição de probabilidade deve refletir o conhecimento que o pesquisador possui sobre o parâmetro em estudo. Para Box e Tiao (1992), dada uma distribuição a priori, que se acredita ser realista, sobre o parâmetro em estudo, então a distribuição posteriori mostra como a informação dos dados (função de verossimilhança) modifica o que se acredita sobre o parâmetro a priori.

Segundo Seber e Wild (1989), uma gama de distribuições a priori são propostas, uma importante classe de distribuições a priori são as não informativas, quando não se têm informações claras sobre os parâmetros. Box e Tiao (1992) discutem que essas distribuições devem ser usadas como um ponto de referência, e com a informação dos dados obter-se a distribuição a posteriori. Para Gelman et al. (2014), justifica-se o uso de prioris não informativas quando pretende-se mostrar que os dados tenham maior peso na informação a posteriori, de modo que as inferências não sejam afetadas por informações externas aos dados.

Outra classe de prioris são as informativas, em que se expressa um maior conhecimento sobre o parâmetro a ser estudado. Segundo Kéry (2010), a escolha de prioris informativas deve ser justificada, pois a escolha da priori pode ter uma influência inadequada na distribuição a posteriori.

Uma importante família de distribuição a priori é a família de distribuições conjugada. Nesse caso, a distribuição a posteriori pertence a mesma família que a distribuição a priori, com novos hiperparâmetros. Segundo Gelman et al. (2014), as distribuições conjugadas são matematicamente convenientes e computacionalmente mais vantajosas já que a distribuição a posteriori tem uma forma conhecida. Gamerman e Lopes (2006) alertam para o cuidado com o uso indiscriminado de prioris conjugadas uma vez que, nem sempre é uma representação adequada da incerteza a priori.

A distribuição a posteriori é fundamental pois toda a inferência em relação aos parâmetros é realizada por meio dessa distribuição.

Ocorre que muitas vezes a distribuição a posteriori é analiticamente intratável. Sendo assim, a obtenção de resumo a posteriori dos parâmetros de interesse, tais como médias e intervalos de credibilidade, são difíceis de serem calculados por envolverem o cálculo de integrais. Uma forma para contornar esse problema está na obtenção de amostras da distribuição a posteriori, pois por meio delas é possível obter resumos a posteriori de todos os parâmetros de interesse.

Amostras da distribuição a posteriori conjunta podem ser obtidas utilizando-se cadeias de Markov que possuam como distribuição limite a distribuição a posteriori, essas cadeias podem ser obtidas utilizando-se o amostrador de Gibbs (CASELLA; GEORGE, 1992) e/ou o algoritmo de Metropolis-Hastings (HASTINGS, 1970). Obtidas as amostras, os resumos a posteriori para cada parâmetro de interesse pode ser obtido utilizando-se o método Monte Carlo daí a origem da sigla MCMC (Markov Chain Monte Carlo). Conforme Gamerman e Lopes (2006), quando se tem a distribuição a posteriori disponível, podem-se resumir informações por meio de alguns elementos. A média da distribuição a posteriori é o valor esperado de  $\theta$  dados os dados, a moda seria o mais provável valor e a mediana divide o espaço paramétrico em duas partes de probabilidades iguais. Ainda para os autores, esses métodos fornecem apenas uma aproximação para a distribuição a posteriori e só devem ser usados quando não é possível extrair informação analiticamente.

#### 2.4.2 Método de Monte Carlo via Cadeia de Markov (MCMC)

Segundo Gamerman e Lopes (2006), uma cadeia de Markov é um processo estocástico em que dado o estado atual da cadeia, o estado passado e futuro são independentes. Para Morettin e Toloi (2004), um processo estocástico  $(Z)$  é definido como: sendo  $T$  um conjunto arbitrário, uma família  $Z = \{Z(t), t \in T\}$  e para cada  $t \in T$ ,  $Z(t)$  é uma variável aleatória.

Os métodos de amostragem Gibbs sampling e Metropolis-Hastings consistem em se obter uma amostra da distribuição conjunta dos parâmetros a partir das distribuições condicionais completas a posteriori e com essa amostra calcular estimativas de média, mediana e quantis.

Segundo Gamerman e Lopes (2006), uma cadeia de Markov é um processo estocástico em que dado o estado atual da cadeia, o estado passado e futuro são independentes. Para Morettin e Toloi (2004), um processo estocástico  $(Z)$  é definido como:

sendo  $T$  um conjunto arbitrário, uma família  $Z = \{Z(t), t \in T\}$  e para cada  $t \in T$ ,  $Z(t)$  é uma variável aleatória. Então, dado um processo estocástico  $\{\phi_0, \phi_1, \dots, \phi_t\}$ , a distribuição de  $\phi_t$  depende apenas de  $\phi_{t-1}$ . Além disso, as primeiras informações das cadeias podem ser retiradas ("Burn in"), visto que valores iniciais não influenciam a distribuição de equilíbrio da cadeia. pois de acordo com o aumento da cadeia, as iterações iniciais são esquecidas.

O método de aproximação tem as vantagens de que sua precisão é controlada pelo tamanho da amostra independentemente do número de observações nos dados e os erros de aproximação podem ser probabilisticamente medidos (GAMERMAN; LOPES, 2006).

### 2.4.3 Amostrador de Gibbs

Inicialmente desenvolvido por Geman e Geman (1984), o amostrador de Gibbs é uma técnica para gerar amostras de uma distribuição conjunta a posteriori de forma indireta sem ser necessário calcular a densidade (CASELLA; GEORGE, 1992), para sua implementação é necessário que as distribuições condicionais completas a posteriori sejam distribuições conhecidas.

Seja a distribuição conjunta posteriori,  $\pi(\boldsymbol{\theta}|\mathbf{X})$ , com  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)'$  e  $p$  sendo número de parâmetros do vetor, para cada um dos parâmetros de  $\pi(\boldsymbol{\theta}|\mathbf{X})$  calcula-se a distribuição condicional completa a posteriori  $\pi(\theta_i|\theta_2, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_p, \mathbf{X})$  que deve ter a forma de uma distribuição já conhecida na literatura.

Conforme Gelfand et al. (1990), Gelfand (2000), Roberts e Smith (1994) e Gamerman e Lopes (2006) considerando-se valores iniciais para os parâmetros  $\boldsymbol{\theta}^{(t=0)} = (\theta_1^{(t=0)}, \theta_2^{(t=0)}, \dots, \theta_p^{(t=0)})'$ , o amostrador de Gibbs é dado da seguinte forma:

- 1) gerar  $\theta_1^{(t+1)}$  a partir de  $\pi(\theta_1|\theta_2^{(t)}, \dots, \theta_p^{(t)}, \mathbf{X})$ ,
- 2) gerar  $\theta_2^{(t+1)}$  a partir de  $\pi(\theta_2|\theta_1^{(t+1)}, \theta_3^{(t)}, \dots, \theta_p^{(t)}, \mathbf{X})$ , ...
- p) gerar  $\theta_p^{(t+1)}$  a partir de  $\pi(\theta_p|\theta_1^{(t+1)}, \theta_2^{(t+1)}, \dots, \theta_{p-1}^{(t+1)}, \mathbf{X})$ ,

ou seja, considerando-se  $t = 0$  obtêm-se  $\boldsymbol{\theta}^{(t=1)} = (\theta_1^{(t=1)}, \theta_2^{(t=1)}, \dots, \theta_p^{(t=1)})'$ . Repetindo-se o processo acima  $T$  vezes, obtêm-se uma amostra de tamanho  $T$  da distribuição conjunta a posteriori. A partir daí pode-se considerar um período de *burn-in*, isto é, queima de valores iniciais e/ou *thin* para diminuir a correlação entre os valores amostrados.

Os valores obtidos no processo do amostrador de Gibbs são uma amostra da distribuição a posteriori. Assim podem-se obter as estimativas pontuais desejadas (média, moda, mediana) e calcular medidas intervalares (intervalo de credibilidade).

Segundo Gelfand (2000), quando a distribuição condicional completa a posteriori,  $\pi(\boldsymbol{\theta}|\mathbf{X})$ , não é uma distribuição conhecida, é necessário utilizar outro método de amostragem. Um método possível é o algoritmo de Metropolis-Hastings.

#### 2.4.4 Metropolis-Hastings

Diferentemente do amostrador de Gibbs, no algoritmo de Metropolis-Hastings as distribuições condicionais completas a posteriori, para um ou mais parâmetros, não apresentam formas conhecidas de distribuição de probabilidade.

O método Metropolis-Hastings proposto por Metropolis et al. (1953) e, posteriormente, generalizado por Hastings (1970) consiste em gerar um valor candidato para um parâmetro e aceitá-lo com uma probabilidade  $\alpha$ .

Segundo Roberts e Smith (1994), o algoritmo constrói uma probabilidade de transição do estado atual,  $X^t = x$ , para o próximo estado realizado  $X^{t+1} = y$  de acordo com uma probabilidade de transição  $\alpha$  definida como:

$$\alpha(x, y) = \begin{cases} \min \left\{ \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}, 1 \right\} & \text{se } \pi(x)q(x, y) > 0 \\ 1 & \text{se } \pi(x)q(x, y) = 0. \end{cases}$$

em que  $q(x, y)$  é a distribuição proposta para o próximo estado da cadeia,  $y$ . O valor gerado é aceito com probabilidade  $\alpha$  e a cadeia permanece no valor atual  $x$  com probabilidade  $1 - \alpha$ .

#### 2.4.5 Convergência das cadeias

Para diagnosticar a convergência das cadeias geradas pelos métodos de amostragem, podem-se utilizar testes para este fim ou verificar de forma gráfica o histórico com a trajetória da cadeia. O primeiro caso tenta medir distâncias e estabelecer limites nas funções de distribuições geradas a partir da cadeia, e a segunda, mais prática, apenas observar a cadeia (GAMERMAN; LOPES, 2006).

Alguns métodos de convergência são propostos na literatura, como os de Gelfand e Smith (1990), Heidelberg e Welch (1983), Gelman e Rubin (1992), Geweke (1992) e Raftery e Lewis (1992). Vários desses métodos são comparados por Nogueira, Sáfiadi e

Ferreira (2004), e os autores propõem uma forma combinada de utilização de vários critérios de convergência de cadeias de Markov.

O método de Gelfand e Smith (1990) é considerado informal, pois se trata de uma análise apenas gráfica em que é verificada a trajetória da cadeia ao longo das iterações. Sendo assim, espera-se que a trajetória alcance a estacionariedade após um certo período percorrido no processo iterativo. Caso ocorra essa estacionariedade, considera-se que houve convergência.

Outro método bastante utilizado com a obtenção de apenas uma cadeia é o de Geweke (1992). O método que também considera a queima de valores iniciais da iteração, consiste em considerar uma função de interesse  $g(\theta)$  e os valores simulados em cada iteração  $g^i = g^1, g^2 \dots$  com  $i = 1, 2, \dots, n$  suficientemente grande, e pode-se considerar a média dessa série temporal  $g^i$ . A partir disso, divide-se a amostra em duas sequências, sendo  $n_A = 0, 1n$  a primeira e  $n_B = 0, 5n$  a segunda. Seguindo o método, calcula-se as médias  $\hat{g}_A$  e  $\hat{g}_B$  e suas respectivas variâncias assintóticas  $\hat{S}_g^A$  e  $\hat{S}_g^B$ . Então, tem-se um Z-score:

$$\frac{\hat{g}_A - \hat{g}_B}{\sqrt{\frac{\hat{S}_g^A}{n_A} + \frac{\hat{S}_g^B}{n_B}}} \sim N(0, 1).$$

Caso a diferença padronizada das médias for grande, há indícios de não convergência da cadeia. O teste ainda pode ser decidido a partir de um valor descritivo (p-valor), considerando um nível de significância  $\alpha$ .

#### 2.4.6 Intervalo de Credibilidade - HPD

Quando se obtém uma estimação pontual de um parâmetro a partir de uma distribuição a posteriori, tem-se toda informação da distribuição resumida apenas em um número. Sendo assim, é importante atribuir um intervalo de credibilidade baseado na distribuição a posteriori, e quanto menor for o tamanho do intervalo, mais concentrada é a distribuição do parâmetro, ou seja, o tamanho do intervalo informa a dispersão sobre o parâmetro (EHLERS, 2007).

Segundo Paulino, Turkman e Murteira (2003), existem diversas regiões de credibilidade com o mesmo grau de credibilidade, porém o interesse está no intervalo de credibilidade de máxima densidade a posteriori HPD (*Highest Posterior Density*) com intervalo de menor comprimento obtido com os valores de maior densidade a posteriori. Con-

forme expresso por Box e Tiao (1992), deve-se ter a propriedade de que a densidade de um ponto dentro do intervalo seja maior do que para pontos fora do intervalo.

Ehlers (2007) define que intervalos de comprimento mínimo são obtidos tomando-se os valores do parâmetro com maior densidade a posteriori. Matematicamente, tem-se que um intervalo de credibilidade  $C$  de  $100(1-\alpha)$  para um parâmetro  $\theta$  é de máxima densidade a posteriori se  $C = \{\theta \in \Theta : \pi(\theta | x) \geq k(\alpha)\}$  sendo  $k(\alpha)$  a maior constante tal que  $P(\theta \in C) \geq 1 - \alpha$ . Ainda para o autor, utilizando esta definição, todos os pontos dentro do intervalo terão densidade maior do que qualquer ponto fora do mesmo.



### 3 MATERIAL E MÉTODOS

#### 3.1 Material

##### 3.1.1 Simulação

Neste trabalho, foi realizado um estudo de simulação para avaliar a influência de diferentes tamanhos amostrais, bem como o efeito da variância residual no processo de estimação dos parâmetros.

O estudo, baseado nas estimativas dos parâmetros disponibilizados em Rezende (2002), foi dividido em seis cenários para cada um dos modelos, apresentados nas Figuras 1 e 2. Os dados simulados representavam o teor de Zn depositado na tibia de frangos, que receberam as dosagens 0, 15, 30, 45, 60, 75, 90, 105 e 120 ppm de Zn na ração.

No primeiro cenário, considerou-se um tamanho amostral de  $n = 16$  indivíduos, no segundo cenário  $n = 30$  indivíduos e no terceiro cenário  $n = 100$  indivíduos. Além disso, nesses três primeiros cenários, foi considerado que o erro segue uma distribuição normal com média 0 e variância constante  $\sigma^2 = 64ppm^2$ , ou seja,  $\epsilon \sim N(0, 64)$ .

No quarto, quinto e sexto cenários, os tamanhos amostrais considerados também foram  $n = 16$ ,  $n = 30$  e  $n = 100$  indivíduos, respectivamente. Porém, foi atribuída uma variância residual para cada uma das nove doses de Zn na ração, isto é, considerou-se  $\epsilon \sim N(0, \sigma_j^2)$ , com  $j=1,2,...,9$ , em que as variâncias de  $\sigma_1^2 = 120,12ppm^2$ ;  $\sigma_2^2 = 66,91ppm^2$ ;  $\sigma_3^2 = 47,33ppm^2$ ;  $\sigma_4^2 = 16,72ppm^2$ ;  $\sigma_5^2 = 46,65ppm^2$ ;  $\sigma_6^2 = 61,31ppm^2$ ;  $\sigma_7^2 = 77,97ppm^2$ ;  $\sigma_8^2 = 61,31ppm^2$  e  $\sigma_9^2 = 96,63ppm^2$  para as dosagens de 0, 15, 30, 45, 60, 75, 90, 105 e 120 ppm de Zn na ração, respectivamente. Esses valores de variâncias residuais foram atribuídos com base na análise realizada por Rezende (2002).



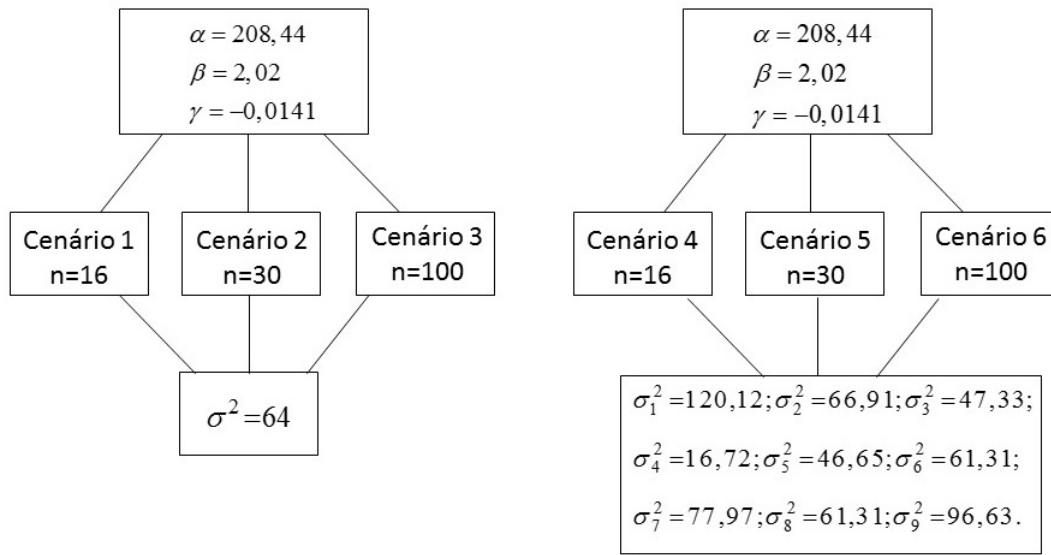


Figura 1 - Valores paramétricos utilizados no estudo de simulação por meio do modelo MQP

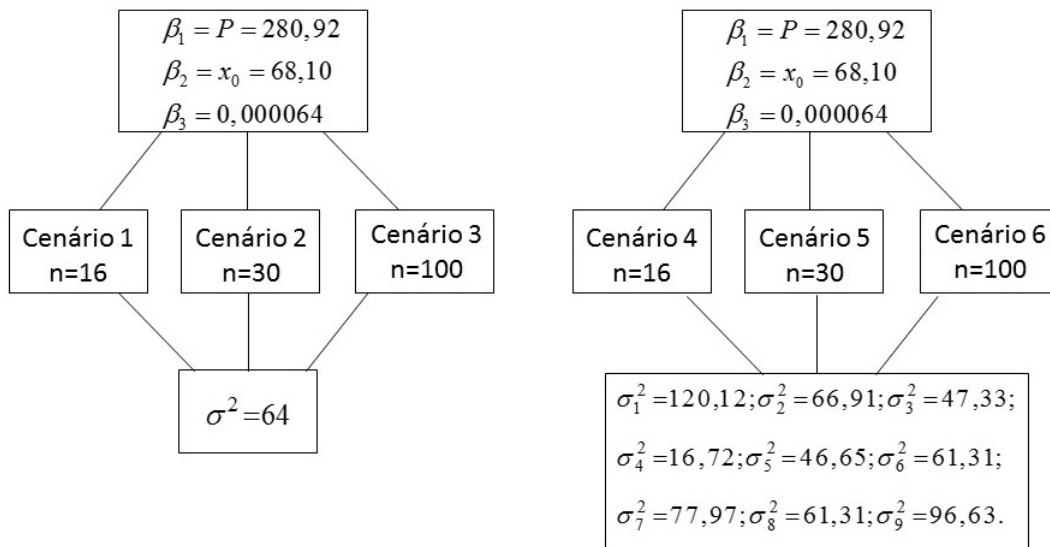


Figura 2 - Valores paramétricos utilizados no estudo de simulação por meio do modelo MNL

Em cada cenário, foram geradas 100 amostras e realizadas 10 mil iterações no processo de estimação bayesiano. Além disso, para desprezar a influência dos valores iniciais dos parâmetros em cada cadeia, foi considerado o descarte (*burn in*) dos primeiros

50% dos valores e, também, foram considerados saltos (*thin*) de 10 valores para diminuir a autocorrelação entre os valores gerados.

Na verificação da adequação das estimativas de cada parâmetro obtidas, foi realizada uma análise exploratória dos cenários simulados por meios gráficos e avaliação do processo com análise do viés.

### 3.1.2 Dados experimentais

Os dados utilizados para o estudo sobre os modelos de regressão com platô de resposta são referentes ao teor de Zinco (Zn), em ppm, na tíbia de frangos e frangas de corte como variável resposta e disponibilizados em Rezende (2002).

O experimento foi instalado no Departamento de Zootecnia e de Solos da Faculdade de Agronomia da Universidade Federal do Rio Grande do Sul (UFRGS), em um delineamento inteiramente ao acaso com 8 repetições, e os tratamentos foram arranjados em esquema fatorial (9x2) sendo 9 dosagens de Zn, em ppm, na ração das aves e o sexo dos animais. As doses de Zn utilizadas foram: 0, 15, 30, 45, 60, 75, 90, 105, 120 ppm, sendo a variável analisada o teor de Zn na tíbia dos frangos.

As aves utilizadas eram da linhagem Hubbard e foram utilizados pintos de corte de um dia de idade, sexados e vacinados contra as doenças de Marek e Bouba Aviária. Foram alojados 12 pintos do mesmo sexo por unidade experimental e, no final da primeira semana, foi feita uma refugagem reduzindo o total para 10 pintos. No 27º dia de idade das aves, foram selecionadas 2 aves aleatoriamente de cada unidade experimental para serem abatidas. Após o abate das aves, as tíbias foram descarnadas, aneladas com anilhas e levadas para o refrigerador para, posteriormente, serem preparadas para as análises, e assim, o teor de Zn foi calculado em relação ao peso da tíbia seca.

Desta forma, tem-se como variável resposta  $y_{ij}$ , com  $i=1,\dots,16$  e  $j=1,\dots,9$ , o teor de Zn em ppm na tíbia da  $i$ -ésima ave que recebeu  $j$ -ésima dose de Zn na ração; e a variável independente  $x_j$  sendo as doses de Zn utilizadas na ração também em ppm.

Para o processo de estimação dos parâmetros, utilizou-se como valores iniciais para os processos de amostragem  $\alpha^0 = 207,4$ ,  $\beta^0 = 1,93$ ,  $\gamma^0 = -0,0129$  e  $\tau^0 = 0,002$ , sendo a precisão  $\tau$  dada pelo inverso da variância, para o modelo de regressão quadrático com platô de resposta, o qual foi utilizado uma variância. Para o modelo não linear com platô de resposta, em que se considerou variância constante entre doses, utilizou-se como valores

iniciais  $\beta_1^0 = 290$ ,  $\beta_2^0 = 70$ ,  $\beta_3^0 = 0,000063$  e  $\tau^0 = 0,004$ . Nos casos em que considerou-se variâncias heterogêneas entre doses, os valores iniciais para o parâmetro de precisão foram  $\tau_1^0 = 0,0005127146$ ,  $\tau_2^0 = 0,0096843088$ ,  $\tau_3^0 = 0,0014566828$ ,  $\tau_4^0 = 0,0011327688$ ,  $\tau_5^0 = 0,0003911969$ ,  $\tau_6^0 = 0,0006729423$ ,  $\tau_7^0 = 0,0039$ ,  $\tau_8^0 = 0,0017$  e  $\tau_9^0 = 0,001653$ . Todos os valores iniciais utilizados, foram baseados em estudos anteriores de Rezende (2002) e Rezende et al. (2007).

As distribuições a priori para os parâmetros do MQP foram distribuição normal como distribuição a priori para  $\alpha$ ,  $\beta$  e  $\gamma$ , e distribuição gama como distribuição a priori para o parâmetro de precisão  $\tau$ . Para os parâmetros do MNL, a distribuição a priori para os parâmetros  $\beta_1$  e  $\beta_2$  foi a distribuição normal, para o parâmetro  $\beta_3$  foi a distribuição beta e para o parâmetro de precisão  $\tau$ , distribuição gama.

Com relação ao número de iterações dos processos de MCMC, foram consideradas 10 mil para os parâmetros do MQP. Já para o modelo MNL, devido a dificuldade de convergência e da autocorrelação dentre as estimativas, foram consideradas 100 mil iterações. Além disso, foi considerado a queima (*burn in*) de 50% das primeiras iterações e saltos (*thin*) de 10 valores para o MQP e de 100 valores para o MNL.

Todas as análises para o presente trabalho foram realizadas no software livre R Core Team (2016). Para o ajuste de modelos na simulação e na análise dos dados reais, foi utilizado o pacote R2OpenBUGS (STURTZ; LIGGES; GELMAN, 2005). Para obtenção do resumo a posteriori para os parâmetros dos modelos e teste de convergência das cadeias, foi utilizado o pacote coda (PLUMMER et al., 2006). Os códigos utilizados no software R estão parcialmente apresentados no anexo A.

## 3.2 Método

### 3.2.1 Modelos empíricos

Os modelos utilizados, tanto no estudo de simulação como no ajuste aos dados experimentais, foram:

Modelo de regressão quadrático com platô de resposta (MQP)

$$y_{ij} = \begin{cases} \alpha + \beta x_j + \gamma x_j^2 + \varepsilon_{ij} & \text{se } x_j \leq x_0; \\ P + \varepsilon_{ij} & \text{se } x_j > x_0. \end{cases}$$

com  $j=1,\dots,9$ , dosagens de Zinco na ração das aves,  $i=1,\dots,n$  observações,  $y_{ij}$  é o teor de

Zn encontrado na tíbia das aves avaliadas em ppm;  $x_j$  são as dosagens de Zn utilizadas na ração em ppm;  $\alpha$ ,  $\beta$ ,  $\gamma$  são parâmetros do modelo;  $x_0$  é o valor da dose de Zn adicionada à ração em que ocorre a estabilização do teor de Zn na tíbia dos frangos (ponto de troca);  $P$  é o platô de resposta, interpretado como o teor máximo de Zn depositado na tíbia dos frangos e  $\varepsilon_{ij}$  é o erro associado ao ajuste do modelo. Os valores para o ponto de troca ( $x_0$ ) e platô de resposta ( $P$ ) foram determinados como funções dos parâmetros da parte quadrática do modelo, ou seja,  $x_0 = -\beta/2\gamma = 71,79$  e  $P = \alpha - (\beta^2/4\gamma) = 281,05$  (Tabela 1).

E o modelo não linear com platô de resposta (MNL)

$$y_{ij} = \begin{cases} \beta_1 \exp(-\beta_3(x_j - \beta_2)^2) + \varepsilon_{ij} & \text{se } x_j \leq x_0; \\ P + \varepsilon_{ij} & \text{se } x_j > x_0. \end{cases}$$

sendo  $j=1,\dots,9$ , dosagens de Zinco na ração das aves,  $i=1,\dots,n$  observações,  $y_{ij}$  é o teor de Zn encontrado na tíbia das aves avaliadas em ppm;  $x_j$  são as dosagens de Zn utilizadas na ração em ppm;  $\beta_1$ ,  $\beta_2$  e  $\beta_3$  são parâmetros do modelo;  $x_0$  é o valor da dose de Zn adicionada à ração em que ocorre a estabilização do teor de Zn na tíbia dos frangos, e que no MNL é igual ao parâmetro  $\beta_2$ ;  $P = \beta_1 \exp(-\beta_3(x_0 - \beta_2)^2)$  é o platô de resposta, e  $\varepsilon_{ij}$  é o erro associado ao ajuste do modelo.

### 3.2.2 Metodologia bayesiana

A estimação dos parâmetros dos modelos MQP e MNL, deu-se por meio da abordagem bayesiana e, para tal, assumiu-se que os erros seguem uma distribuição normal. Desta forma, a função de verossimilhança para o modelo MQP, pode ser escrita como:

$$\begin{aligned}
L(\boldsymbol{\theta}|y) &= \prod_{i=1}^{16} \prod_{j=1}^k \left( \frac{\tau}{2\pi} \right)^{\frac{1}{2}} \exp \left\{ \frac{-\tau}{2} (y_{ij} - (\alpha + \beta x_j + \gamma x_j^2))^2 \right\} \\
&\times \prod_{i=1}^{16} \prod_{j=k+1}^9 \left( \frac{\tau}{2\pi} \right)^{\frac{1}{2}} \exp \left\{ -\frac{\tau}{2} (y_{ij} - P)^2 \right\} \\
&= \left( \frac{\tau}{2\pi} \right)^{\frac{k}{2}} \exp \left\{ \frac{-\tau}{2} \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\alpha + \beta x_j + \gamma x_j^2))^2 \right\} \\
&\times \left( \frac{\tau}{2\pi} \right)^{\frac{n-k}{2}} \exp \left\{ \frac{-\tau}{2} \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right\} \\
&= \left( \frac{\tau}{2\pi} \right)^{\frac{n}{2}} \exp \left\{ \frac{-\tau}{2} \left[ \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\alpha + \beta x_j + \gamma x_j^2))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right] \right\}
\end{aligned}$$

com  $\boldsymbol{\theta}' = (\alpha, \beta, \gamma, \tau)'$  o vetor de parâmetros do modelo, e k sendo a posição da maior dosagem não superior a  $x_0$ .

Para os parâmetros  $\alpha$ ,  $\beta$ ,  $\gamma$  foram atribuídas distribuições normais a priori, ou seja:

Para o parâmetro  $\alpha \sim N(\mu_\alpha, \tau_\alpha)$ :

$$\pi(\alpha) = \left( \frac{\tau_\alpha}{2\pi} \right)^{\frac{1}{2}} \exp \left\{ \frac{-\tau_\alpha}{2} (\alpha - \mu_\alpha)^2 \right\}.$$

Para o parâmetro  $\beta \sim N(\mu_\beta, \tau_\beta)$ , tem-se que:

$$\pi(\beta) = \left( \frac{\tau_\beta}{2\pi} \right)^{\frac{1}{2}} \exp \left\{ \frac{-\tau_\beta}{2} (\beta - \mu_\beta)^2 \right\}.$$

Já para o parâmetro  $\gamma \sim N(\mu_\gamma, \tau_\gamma)$ , a distribuição a priori fica:

$$\pi(\gamma) = \left( \frac{\tau_\gamma}{2\pi} \right)^{\frac{1}{2}} \exp \left\{ \frac{-\tau_\gamma}{2} (\gamma - \mu_\gamma)^2 \right\}.$$

Para o parâmetro da precisão  $\tau$ , foi considerada uma distribuição a priori

gama, ou seja,  $\tau \sim \text{gama}(a_\tau, b_\tau)$ :

$$\pi(\tau) = \tau^{a_\tau-1} e^{-\tau b_\tau}.$$

De posse das distribuições a priori de cada parâmetro e da função de verossimilhança para o modelo MQP, é possível obter a distribuição a posteriori conjunta, dada por:

$$\pi(\alpha, \beta, \gamma, \tau | y, H) \propto L(\boldsymbol{\theta} | y) \pi(\alpha | \mu_\alpha, \tau_\alpha) \pi(\beta | \mu_\beta, \tau_\beta) \pi(\gamma | \mu_\gamma, \tau_\gamma) \pi(\tau | a_\tau, b_\tau)$$

sendo  $H = \{\mu_\alpha, \tau_\alpha, \mu_\beta, \tau_\beta, \mu_\gamma, \tau_\gamma, a_\tau, b_\tau\}$  o conjunto dos parâmetros das distribuições a priori denominados de hiperparâmetros. Os valores considerados para os hiperparâmetros foram obtidos por meio de estudos anteriores em que foram utilizadas estimações por métodos clássicos para obtenção de suas estimativas.

Considerando a distribuição a posteriori conjunta para os parâmetros do MQP, é possível encontrar as distribuições a posteriori condicionais para cada parâmetro, sendo essas utilizadas nos métodos de amostrador de Gibbs ou algoritmo de Metropolis-Hastings para simular valores da distribuição de cada parâmetro.

A distribuição a posteriori condicional para o parâmetro  $\alpha$  do MQP é dada por:

$$\pi(\alpha | y, \beta, \gamma, \tau) = \exp \left\{ \frac{-\tau}{2} \left[ \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\alpha + \beta x_j + \gamma x_j^2))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right] - \frac{\tau_\alpha}{2} (\alpha - \mu_\alpha)^2 \right\}$$

Para o parâmetro  $\beta$  do MQP a distribuição a posteriori condicional é:

$$\pi(\beta | y, \alpha, \gamma, \tau) = \exp \left\{ \frac{-\tau}{2} \left[ \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\alpha + \beta x_j + \gamma x_j^2))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right] - \frac{\tau_\beta}{2} (\beta - \mu_\beta)^2 \right\}$$

O parâmetro  $\gamma$  para o MQP obteve como distribuição a posteriori condicional a seguinte expressão:

$$\pi(\gamma | y, \alpha, \beta, \tau) = \exp \left\{ \frac{-\tau}{2} \left[ \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\alpha + \beta x_j + \gamma x_j^2))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right] - \frac{\tau_\gamma}{2} (\gamma - \mu_\gamma)^2 \right\}$$

e por fim, para o parâmetro de precisão  $\tau$  do MQP, a distribuição a posteriori condicional

foi:

$$\begin{aligned}
\pi(\tau|y, \alpha, \beta, \gamma) &= \tau^{\frac{N}{2}} \exp \left\{ \frac{-\tau}{2} \left[ \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\alpha + \beta x_j + \gamma x_j^2))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_j - P)^2 \right] \right\} \\
&\times \tau^{a_\tau - 1} \exp(-\tau b_\tau) \\
&= \tau^{(\frac{N}{2} + a_\tau) - 1} \\
&\times \exp \left\{ - \left[ \frac{1}{2} \left( \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\alpha + \beta x_j + \gamma x_j^2))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right) + b_\tau \right] \tau \right\}
\end{aligned}$$

o que caracteriza o núcleo de uma distribuição gama com parâmetros:

$$a^* = \frac{n}{2} + a_\tau$$

e

$$b^* = b_\tau + \frac{1}{2} \left( \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\alpha + \beta x_j + \gamma x_j^2))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right)$$

Para os parâmetros  $\alpha$ ,  $\beta$  e  $\gamma$  do MQP, as distribuições a posteriori condicionais não apresentaram formas de distribuições já conhecidas, por isso o método de Metropolis-Hastings é adequado para obtenção da amostra dos valores de tais parâmetros. Já o parâmetro de precisão  $\tau$  apresentou uma distribuição a posteriori condicional que é um núcleo de uma distribuição gama, ou seja, é adequado a utilização do amostrador de Gibbs para se obter a amostra dos valores do parâmetro.

Para o modelo não linear com platô de resposta (MNL), também foi considerado que os erros seguem uma distribuição normal e, portanto, a função de verossimilhança

pode ser escrita como:

$$\begin{aligned}
L(\boldsymbol{\theta}|y) &= \prod_{i=1}^{16} \prod_{j=1}^k \left( \frac{\tau}{2\pi} \right)^{\frac{1}{2}} \exp \left\{ \frac{-\tau}{2} (y_{ij} - (\beta_1 \exp\{-\beta_3(x_j - \beta_2)^2\}))^2 \right\} \\
&\times \prod_{i=1}^{16} \prod_{k+1}^9 \left( \frac{\tau}{2\pi} \right)^{\frac{1}{2}} \exp \left\{ -\frac{\tau}{2} (y_{ij} - P)^2 \right\} \\
&= \left( \frac{\tau}{2\pi} \right)^{\frac{k}{2}} \exp \left\{ \frac{-\tau}{2} \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\beta_1 \exp\{-\beta_3(x_j - \beta_2)^2\}))^2 \right\} \\
&\times \left( \frac{\tau}{2\pi} \right)^{\frac{n-k}{2}} \exp \left\{ \frac{-\tau}{2} \sum_{i=1}^{16} \sum_{k+1}^9 (y_{ij} - P)^2 \right\} \\
&= \left( \frac{\tau}{2\pi} \right)^{\frac{n}{2}} \exp \left\{ \frac{-\tau}{2} \left[ \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\beta_1 \exp\{-\beta_3(x_j - \beta_2)^2\}))^2 + \sum_{i=1}^{16} \sum_{k+1}^9 (y_{ij} - P)^2 \right] \right\}
\end{aligned}$$

com  $\boldsymbol{\theta} = (\beta_1, \beta_2, \beta_3, \tau)'$  o vetor de parâmetros do MNL, e k sendo a posição da maior dosagem não superior a  $x_0$ .

As distribuições a priori consideradas para cada parâmetro do MNL foram para o parâmetro  $\beta_1 \sim N(\mu_{\beta_1}, \tau_{\beta_1})$ :

$$\pi(\beta_1) = \left( \frac{\tau_{\beta_1}}{2\pi} \right)^{\frac{1}{2}} \exp \left\{ \frac{-\tau_{\beta_1}}{2} (\beta_1 - \mu_{\beta_1})^2 \right\}$$

Já para o parâmetro  $\beta_2 \sim N(\mu_{\beta_2}, \tau_{\beta_2})$ , tem-se que a distribuição a priori é:

$$\pi(\beta_2) = \left( \frac{\tau_{\beta_2}}{2\pi} \right)^{\frac{1}{2}} \exp \left\{ \frac{-\tau_{\beta_2}}{2} (\beta_2 - \mu_{\beta_2})^2 \right\}$$

O parâmetro  $\beta_3$  teve como distribuição a priori  $\beta_3 \sim \text{beta}(\alpha_{\beta_3}, \beta_{\beta_3})$ , ou seja:

$$\pi(\beta_3) = \frac{1}{B(\alpha_{\beta_3}, \beta_{\beta_3})} \beta_3^{\alpha_{\beta_3}-1} (1 - \beta_3)^{\beta_{\beta_3}-1}$$

E para o parâmetro de precisão  $\tau \sim \text{gama}(a_\tau, b_\tau)$ :

$$\pi(\tau) = \tau^{a_\tau-1} e^{-\tau b_\tau}$$

com  $\mu_\alpha, \tau_\alpha, \mu_\beta, \tau_\beta, \alpha_\gamma, \beta_\gamma, a_\tau, b_\tau$  os hiperparâmetros das distribuições a priori para os parâmetros do MNL.



Após combinar as distribuições a priori dos parâmetros e função de verossimilhança, foram obtidas as seguintes distribuições condicionais para os parâmetros: Para  $\beta_1$ ,

$$\pi(\beta_1|y, \beta_2, \beta_3, \tau) = \exp \left\{ \frac{-\tau}{2} \left[ \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\beta_1 \exp\{-\beta_3(x_j - \beta_2)^2\}))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right] - \frac{\tau\beta_1}{2}(\beta_1 - \mu_{\beta_1})^2 \right\}$$

Para  $\beta_2$ ,

$$\pi(\beta_2|y, \beta_1, \beta_3, \tau) = \exp \left\{ \frac{-\tau}{2} \left[ \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\beta_1 \exp\{-\beta_3(x_j - \beta_2)^2\}))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right] - \frac{\tau\beta_2}{2}(\beta_2 - \mu_{\beta_2})^2 \right\}$$

Para  $\beta_3$ ,

$$\begin{aligned} \pi(\beta_3|y, \beta_1, \beta_2, \tau) &= \exp \left\{ \frac{-\tau}{2} \left[ \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\beta_1 \exp\{-\beta_3(x_j - \beta_2)^2\}))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right] \right\} \\ &\times \beta_3^{\alpha_{\beta_3}-1} (1 - \beta_3)^{\beta_{\beta_3}-1} \end{aligned}$$

e para  $\tau$  do MNL,

$$\begin{aligned} \pi(\tau|y, \beta_1, \beta_2, \beta_3) &= \left( \frac{\tau}{2} \right)^{\frac{N}{2}} \\ &\times \exp \left\{ \frac{-\tau}{2} \left[ \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\beta_1 \exp\{-\beta_3(x_j - \beta_2)^2\}))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right] \right\} \\ &\times \tau^{a_\tau-1} e^{-\tau b_\tau} \\ &= \tau^{\left(\frac{N}{2} + a_\tau\right) - 1} \\ &\times \exp \left\{ - \left[ \frac{1}{2} \left( \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\beta_1 \exp\{-\beta_3(x_j - \beta_2)^2\}))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right) + b_\tau \right] \tau \right\} \end{aligned}$$

A expressão da distribuição a posteriori condicional do parâmetro  $\tau$  tem forma conhecida é um núcleo de uma distribuição gama com parâmetros  $\alpha^*$  e  $\beta^*$ , sendo:

$$\alpha^* = \frac{n}{2} + a_\tau$$

e

$$\beta^* = b_\tau + \frac{1}{2} \left( \sum_{i=1}^{16} \sum_{j=1}^k (y_{ij} - (\beta_1 \exp\{-\beta_3(x_j - \beta_2)^2\}))^2 + \sum_{i=1}^{16} \sum_{j=k+1}^9 (y_{ij} - P)^2 \right)$$

Considerando as distribuições a posteriori condicionais para cada parâmetro do MNL tem-se que os parâmetros  $\beta_1$ ,  $\beta_2$  e  $\beta_3$  não apresentam forma de alguma distribuição

conhecida, desta forma, o método de Metropolis-Hastings é adequado para gerar valores da distribuição a posteriori de cada parâmetro. Para o parâmetro  $\tau$ , a sua distribuição a posteriori condicional apresenta um núcleo de uma gama, o que permite a utilização do amostrador de Gibbs para simular valores da distribuição a posteriori condicional.

### 3.2.3 Diagnóstico e critério de comparação dos ajustes

Para avaliação dos resíduos dos ajustes dos modelos utilizados, foram verificados os gráficos de resíduos versus valores preditos. Além disso, foi utilizado como critério para escolha do melhor modelo que se ajusta aos dados o critério de informação deviance (DIC), proposto por Spiegelhalter et al. (2002), dado por:

$$\begin{aligned} \text{DIC} &= \hat{D} + 2p_D \\ &= \bar{D} + p_D, \end{aligned}$$

com  $p_D = \bar{D} - \hat{D}$ ,  $\bar{D}$  sendo a média a posteriori da deviance e  $\hat{D} = -2 \times \log(p(y|\bar{\theta}))$ . Sendo assim, o melhor modelo, segundo o DIC, é aquele que apresentar menor valor deste critério.



## 4 RESULTADOS E DISCUSSÃO

### 4.1 Estudo de simulação para o modelo de regressão quadrático com platô de resposta (MQP)

A seguir serão apresentados os resultados do estudo de simulação por meio do modelo de regressão quadrático com platô de resposta (MQP) considerando variância constante,  $\sigma^2=64$  ( $\tau = 1/\sigma^2 = 0,0156$ ), diferentes tamanhos amostrais ( $n=16, 30$  e  $100$ ) e os seguintes valores paramétricos  $\alpha = 208,44$ ,  $\beta = 2,02$ ,  $\gamma = -0,0141$ . Foram simulados 100 experimentos e, para cada um deles, procedeu-se o ajuste do modelo MQP por meio da metodologia bayesiana.

Tabela 1 - Médias das estimativas da média a posteriori dos parâmetros do modelo de regressão quadrático com platô de resposta e vieses médios, para os cenários 1, 2 e 3.

Parâmetros	$\alpha$	$\beta$	$\gamma$	$x_0$	P	$\sigma^2$
Valor paramétrico	208,44	2,02	-0,0141	71,79	281,05	64,00
n=16 (Cenário 1)	208,38 (-0,07)	2,06 (0,03)	-0,0146 (-0,0005)	71,10 (-0,68)	280,91 (-0,01)	65,48 (1,48)
n=30 (Cenário 2)	208,14 (-0,31)	2,06 (0,03)	-0,0145 (-0,0004)	71,34 (-0,45)	281,24 (0,19)	64,53 (0,53)
n=100 (Cenário 3)	208,26 (-0,18)	2,05 (0,03)	-0,0144 (-0,0003)	71,28 (-0,51)	281,29 (0,24)	64,46 (0,46)

Ainda para o modelo de regressão quadrático com platô de resposta (MQP), outros três cenários de simulação (Tabelas 2 e 3) foram considerados utilizando os mesmos valores paramétricos descritos anteriormente, porém considerando variâncias heterogêneas em relação às doses de Zn na ração no estudo de simulação. As variâncias consideradas foram  $\sigma_1^2 = 120,12$ ;  $\sigma_2^2 = 66,91$ ;  $\sigma_3^2 = 47,33$ ;  $\sigma_4^2 = 16,72$ ;  $\sigma_5^2 = 46,65$ ;  $\sigma_6^2 = 61,31$ ;  $\sigma_7^2 = 77,97$ ;  $\sigma_8^2 = 61,31$  e  $\sigma_9^2 = 96,63$ .

Tabela 2 - Médias das estimativas da média a posteriori dos parâmetros do modelo de regressão quadrático com platô de resposta e vieses médios, para os cenários 4, 5 e 6.

Parâmetro	$\alpha$	$\beta$	$\gamma$	$x_0$	P
Valor paramétrico	208,44	2,02	-0,0141	71,79	281,05
n=16	208,35	2,05	-0,0144	71,42	281,10
(Cenário 4)	(-0,09)	(0,02)	(-0,0003)	(-0,36)	(0,05)
n=30	207,74	2,07	-0,0145	71,44	281,29
(Cenário 5)	(-0,71)	(0,04)	(-0,0004)	(-0,35)	(0,24)
n=100	208,22	2,04	-0,0142	71,83	281,42
(Cenário 6)	(-0,22)	(0,02)	(-0,0001)	(0,05)	(0,37)

Tabela 3 - Médias das estimativas da média a posteriori dos parâmetros de variância do modelo de regressão quadrático com platô de resposta e vieses médios, nos cenários 4, 5 e 6.

Parâmetro	$\sigma_1^2$	$\sigma_2^2$	$\sigma_3^2$	$\sigma_4^2$	$\sigma_5^2$	$\sigma_6^2$	$\sigma_7^2$	$\sigma_8^2$	$\sigma_9^2$
Valor paramétrico	120,12	66,91	47,33	16,73	46,65	61,31	77,97	61,31	96,63
n=16	135,61	79,83	55,51	19,18	55,98	68,87	85,19	75,12	113,41
(Cenário 4)	(15,49)	(12,92)	(8,18)	(2,45)	(9,33)	(7,56)	(7,22)	(13,81)	(16,78)
n=30	132,15	71,88	50,85	18,16	53,85	62,59	82,71	68,09	104,23
(Cenário 5)	(12,03)	(4,97)	(3,51)	(1,44)	(7,20)	(1,29)	(4,74)	(6,78)	(7,60)
n=100	120,92	67,23	48,91	17,07	50,70	61,15	79,10	62,54	98,22
(Cenário 6)	(0,80)	(0,32)	(1,58)	(0,35)	(4,05)	(-0,16)	(1,13)	(1,23)	(1,59)

Nas Tabelas 1, 2 e 3 estão apresentadas as médias das estimativas das médias a posteriori nos 100 experimentos, bem como a média dos vieses em relação aos valores paramétricos para cada um dos seis cenários de simulação. Em relação às médias das estimativas é possível perceber, para todos os parâmetros da parte sistemática do modelo, que elas estão bem próximas dos valores paramétricos considerados, tanto para os cenários em que considerou-se variância constante, como para os cenários em que assumiu-se uma variância para cada nível do fator dose.

Na inspeção dos vieses, também é possível verificar (Tabela 1) que ao considerar variância constante, os parâmetros  $\alpha$ ,  $\gamma$  e  $x_0$  tendem a ser subestimados, e esse mesmo comportamento é verificado para os parâmetros  $\alpha$  e  $\gamma$  ao considerar variâncias heterogêneas

(Tabela 2) e para o parâmetro  $x_0$  quando foram considerados os tamanhos amostrais  $n=16$  e  $n=30$ .

Em relação ao parâmetro da variância, ele foi superestimado em todos os cenários considerados. Além disso, o viés médio diminuiu consideravelmente com o tamanho amostral. Para o parâmetro de platô  $P$ , de acordo com as Tabelas 1 e 2, o viés médio é maior para amostras pequenas, porém, na Figura 4, observa-se que a dispersão dos vieses é menor com aumento do tamanho amostral, ou seja, há maior precisão com aumento do tamanho amostral.

Nas Figuras 3 e 4, estão representadas por meio do boxplot as estimativas da média a posteriori dos parâmetros do MQP e vieses em cada um dos cenários de simulação. Observa-se que com o aumento do tamanho amostral a dispersão das estimativas e dos vieses diminuem consideravelmente, tanto nos cenários em que considerou-se variância constante (cenários 1, 2 e 3), como nos casos de heterogeneidade de variâncias entre os níveis do fator dose (cenário 4, 5 e 6).

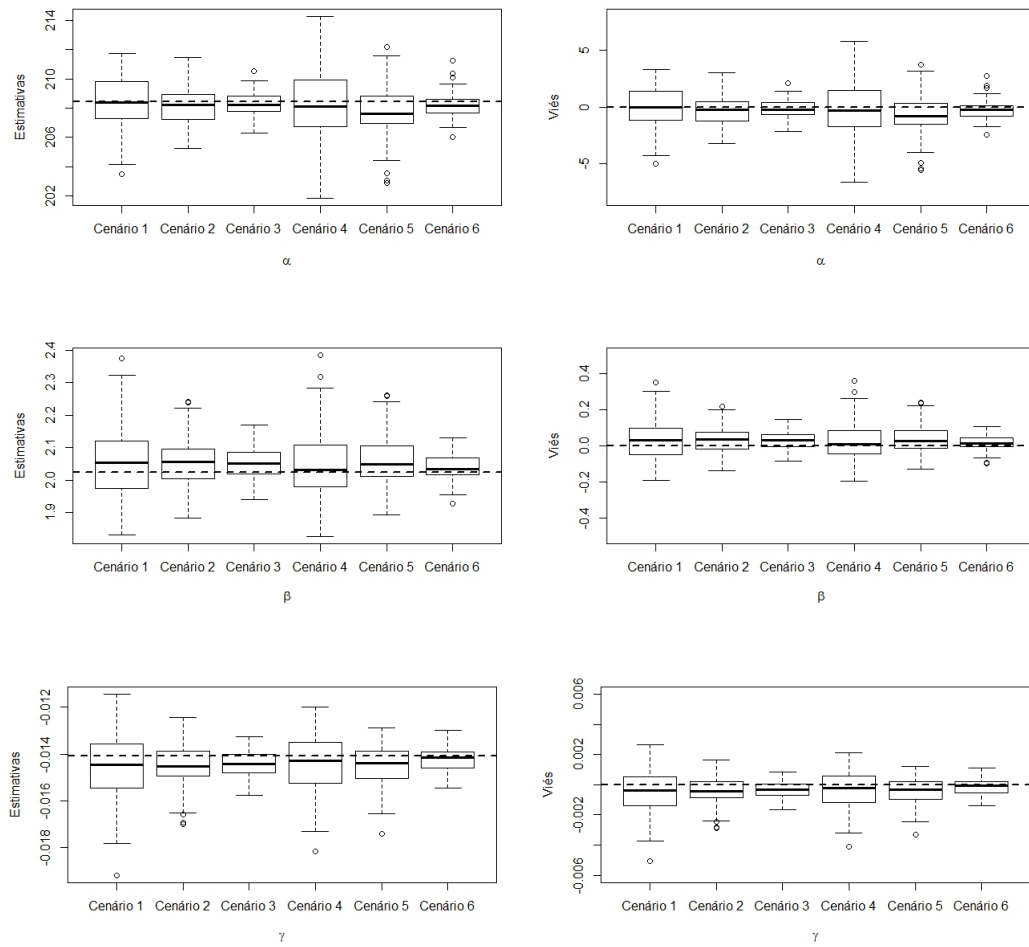


Figura 3 - Boxplot das estimativas e vieses dos parâmetros  $\alpha$ ,  $\beta$  e  $\gamma$  do modelo de regressão quadrático com platô de resposta em cada cenário simulado

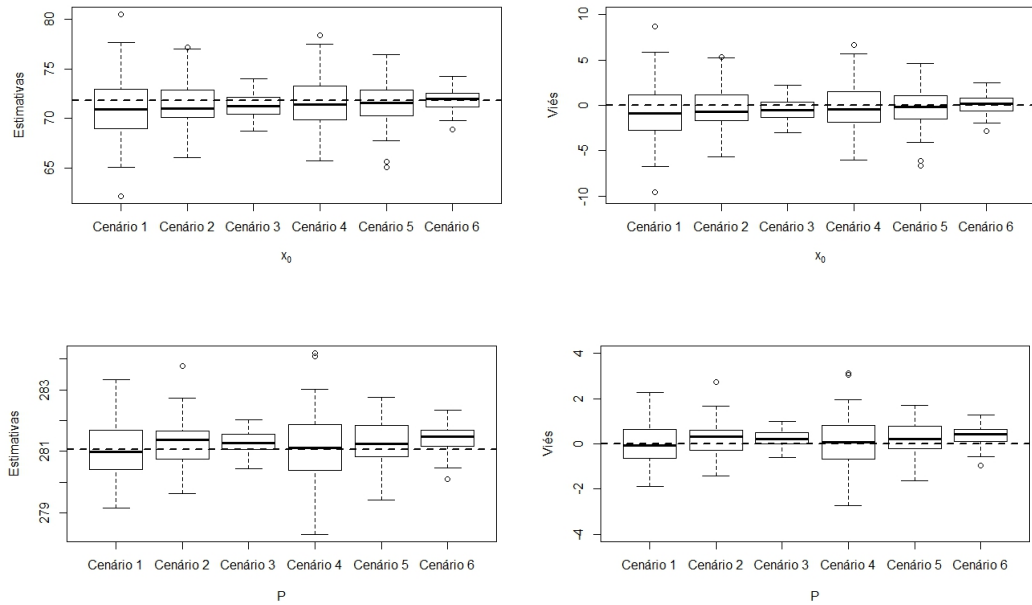


Figura 4 - Boxplot das estimativas e vieses dos parâmetros  $x_0$  e  $P$  do modelo de regressão quadrático com platô de resposta em cada cenário simulado

## 4.2 Estudo de simulação para o modelo não linear com platô de resposta (MNL)

Nesta seção, serão apresentados os resultados do estudo de simulação para o modelo não linear com platô de resposta (MNL), também considerando variância constante,  $\sigma^2=64$  ( $\tau = 1/\sigma^2 = 0,0156$ ), diferentes tamanhos amostrais ( $n=16, 30$  e  $100$ ) e os seguintes valores paramétricos  $\beta_1 = 280,92$ ,  $\beta_2 = 68,10$ ,  $\beta_3 = 0,000064$ . Também para este modelo, foram consideradas variâncias heterogêneas entre os níveis do fator dose ( $\sigma_1^2 = 120,12$ ;  $\sigma_2^2 = 66,91$ ;  $\sigma_3^2 = 47,33$ ;  $\sigma_4^2 = 16,72$ ;  $\sigma_5^2 = 46,65$ ;  $\sigma_6^2 = 61,31$ ;  $\sigma_7^2 = 77,97$ ;  $\sigma_8^2 = 61,31$  e  $\sigma_9^2 = 96,63$ ). Foram simulados 100 experimentos e para cada um deles, procedeu-se o ajuste do modelo MNL por meio da metodologia bayesiana.

Nas Tabelas 4, 5 e 6 estão apresentados os resultados do estudo de simulação para o MNL, considerando os cenários 1, 2, e 3 (Tabela 4) e cenários 4, 5 e 6 (Tabelas 5 e 6). Para o parâmetro  $\beta_1$ , que neste modelo é o próprio platô de resposta, em pequenas amostras ( $n=16$ ), é possível verificar que as médias das médias a posteriori são diferentes entre si, quando a variância foi considerada homogênea (Tabela 4) ou heterogênea (Tabela 5). À



medida que o tamanho amostral aumenta, essa diferença entre estimativas diminui. Esse mesmo comportamento é verificado para o parâmetro  $\beta_2$  que para o modelo MNL é o ponto de troca. O viés é menor nos casos em que foi considerada variância constante para dose, para o parâmetro do ponto de troca  $\beta_2$ . Esse mesmo comportamento pode ser observado no gráfico de boxplot na Figura 5, que apresenta as estimativas da média a posteriori dos parâmetros do MQP e vieses em cada um dos cenários de simulação.

Para o parâmetro de variância (Tabelas 4 e 6), em todos os cenários considerados, sua estimativa foi superestimada, embora o viés médio tenha diminuído consideravelmente com o aumento do tamanho amostral.

Tabela 4 - Médias das estimativas da média a posteriori dos parâmetros do modelo não linear com platô de resposta e vieses médios, para os cenários 1, 2 e 3.

Parâmetro	$\beta_1 = P$	$\beta_2 = x_0$	$\beta_3$	$\sigma^2$
Valor paramétrico	280,92	68,10	0,000064	64
n=16	282,04	68,49	0,000064	67,01
(Cenário 1)	(1,12)	(0,39)	(-0,0000001)	(3,01)
n=30	281,07	68,25	0,000064	64,88
(Cenário 2)	(0,15)	(0,15)	(-0,0000002)	(0,88)
n=100	281,06	68,25	0,000064	64,43
(Cenário 3)	(0,14)	(0,15)	(0,00000008)	(0,43)

Tabela 5 - Médias das estimativas da média a posteriori dos parâmetros do modelo não linear com platô de resposta e vieses médios, para os cenários 4, 5 e 6.

Parâmetro	$\beta_1 = P$	$\beta_2 = x_0$	$\beta_3$
Valor paramétrico	280,92	68,10	0,000064
n=16	280,67	66,03	0,000069
(Cenário 4)	(-0,25)	(-2,07)	(0,000006)
n=30	281,34	69,02	0,000062
(Cenário 5)	(0,43)	(0,92)	(-0,000002)
n=100	281,31	68,56	0,000063
(Cenário 6)	(0,39)	(0,46)	(-0,0000004)

Tabela 6 - Médias das estimativas da média a posteriori dos parâmetros de variância do modelo não linear com platô de resposta e vieses médios, nos cenários 4, 5 e 6.

Parâmetro	$\sigma_1^2$	$\sigma_2^2$	$\sigma_3^2$	$\sigma_4^2$	$\sigma_5^2$	$\sigma_6^2$	$\sigma_7^2$	$\sigma_8^2$	$\sigma_9^2$
Valor paramétrico	120,12	66,91	47,33	16,73	46,65	61,31	77,97	61,31	96,63
n=16	137,32	78,37	52,14	18,66	53,42	70,56	91,94	67,87	111,73
(Cenário 4)	(17,20)	(11,45)	(4,81)	(1,93)	(6,77)	(9,26)	(13,97)	(6,56)	(15,10)
n=30	131,13	73,02	50,10	17,49	50,62	69,27	84,49	67,65	106,06
(Cenário 5)	(11,01)	(6,11)	(2,76)	(0,77)	(3,97)	(7,96)	(6,52)	(6,34)	(9,43)
n=100	123,71	69,66	48,91	17,30	48,82	61,75	80,18	63,12	98,26
(Cenário 6)	(3,58)	(2,75)	(1,58)	(0,57)	(2,17)	(0,44)	(2,21)	(1,81)	(1,63)

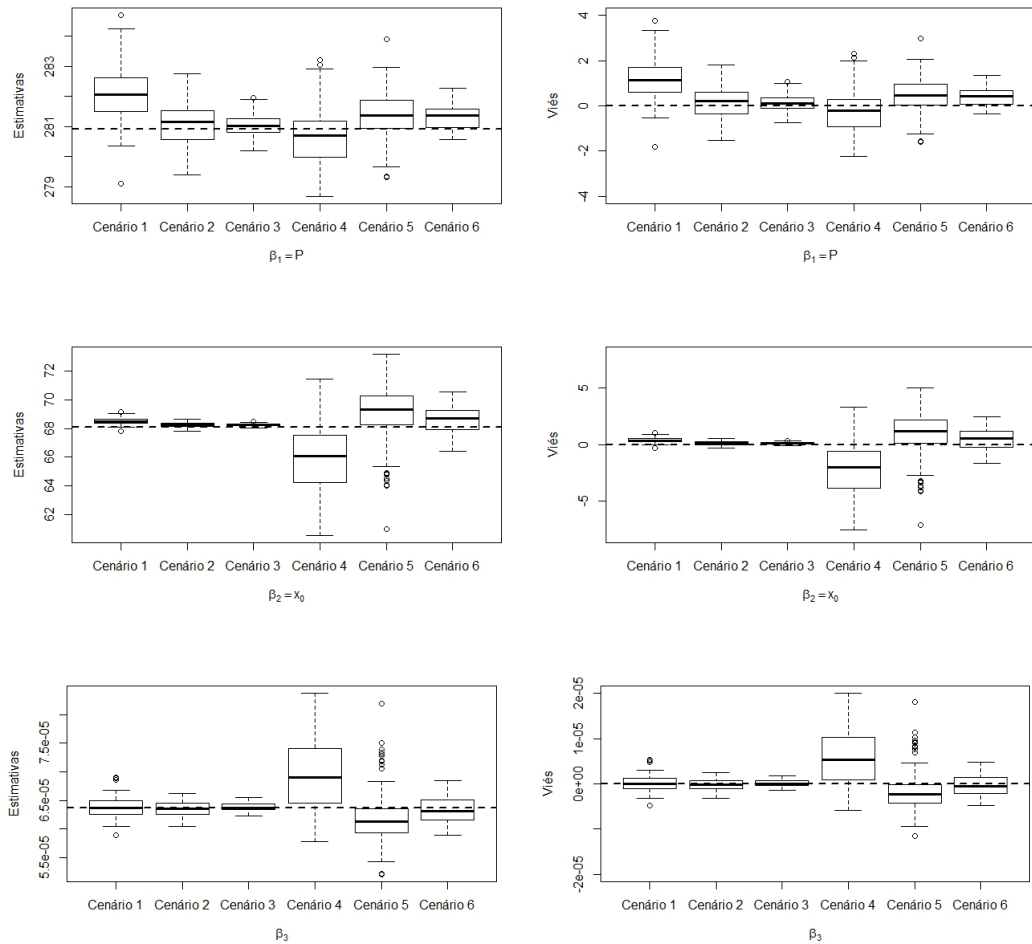


Figura 5 - Estimativas dos parâmetros do modelo não linear com platô de resposta em cada cenário simulado

Além das análises apresentadas para os modelos MQP e MNL, segue no anexo B os histogramas das densidades a posteriori para cada parâmetro dos modelos, gráficos das estimativas a posteriori em função do tamanho amostral e também tabelas com as estimativas dos parâmetros de precisão nos casos de heterogeneidade de variâncias.

### 4.3 Análise dos dados de teor de Zinco (Zn) depositados na tíbia de frangos

Os dados do teor de Zinco (Zn) depositado na tíbia das aves em função das doses de Zinco (Zn) aplicadas na ração estão representados na Figura 6, por meio do box-plot. A estrutura apresentada pelos dados mostra que há um crescimento no teor de Zn encontrado na tíbia das aves, como uma função do aumento da suplementação de Zn considerada na ração. Conforme é visto na Figura 6, há seis valores de teor de Zn na tíbia, que podem ser considerados discrepantes em relação aos demais. Sendo assim, optou-se por realizar o ajuste dos modelos retirando esses pontos.

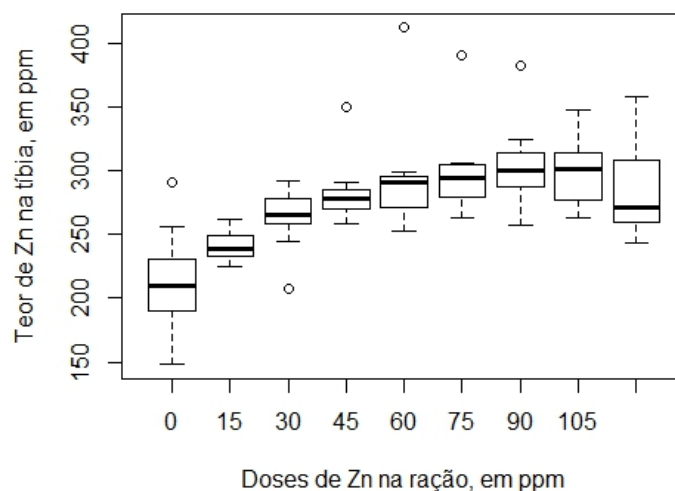


Figura 6 - Boxplot para teor de Zn na tíbia de frangos de corte.

O comportamento dos dados sem os pontos discrepantes pode ser visualizado na Figura 7. É possível perceber que com o aumento da suplementação do nutriente na ração há um aumento na deposição do mesmo na tíbia dos animais. Esse aumento de teor de Zn na tíbia é mais acentuado até, aproximadamente, 75 ppm de Zn na ração, e a partir

desta dose, aparentemente, a deposição deste nutriente tende a se estabilizar.

Esse comportamento está de acordo com o descrito por Euclydes e Rostagno (2001), que caracterizam o efeito de um nutriente limitante na ração partindo de níveis baixos até níveis elevados, em quatro fases. Na primeira fase, denominada inicial, o acréscimo de nutrientes garante apenas a sobrevivência do animal, pois os níveis são insuficientes para ocorrer crescimento. Na segunda fase ou fase resposta, o animal começa a apresentar crescimento e melhor eficiência alimentar até um nível de estabilização do desempenho e início da terceira fase denominada estável. Na fase estável, os níveis de nutrientes não promovem resposta ao desempenho. E, por fim, a quarta fase chamada de tóxica, em que o nível elevado do nutriente pode causar redução no desempenho em consequência de efeitos colaterais. Sendo assim, os autores afirmam que os experimentos para determinar níveis ótimos dos nutrientes na ração devem ser aplicados nas fases de resposta e estável.

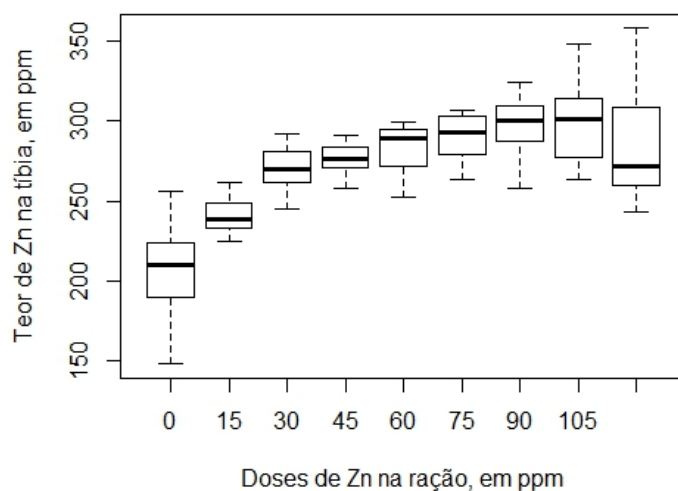


Figura 7 - Boxplot para teor de Zn na tíbia de frangos de corte, sem os outliers.

As médias do teor de Zn na tíbia dos animais de cada dose de Zn aplicada na ração estão apresentadas na Tabela 7, juntamente com os respectivos desvios padrões. Pode-se verificar que o desvio padrão do teor Zn é bastante diferente em cada nível do fator dose, ou seja, pode-se afirmar que há uma heterogeneidade de variâncias. Mesmo assim, procedeu-se com o ajuste do modelos com platô de resposta considerando tanto a homogeneidade de variância como a heterogeneidade da mesma.

Tabela 7 - Médias e desvios padrões do teor de Zn na tíbia dos frangos em cada dose de Zn.

Doses	0	15	30	45	60	75	90	105	120
Média	207,13	241,20	270,27	275,81	282,31	289,35	295,87	298,61	284,66
Desvio padrão	28,73	10,85	14,46	10,18	15,16	14,21	19,07	25,97	35,17

#### 4.3.1 Ajuste do modelo de regressão quadrático com platô de resposta (MQP) e do modelo não linear com platô de resposta(MNL), considerando variância homogênea

O processo de estimação dos parâmetros procedeu-se por meio da abordagem bayesiana utilizando métodos numéricos para obtenção de amostras das distribuições a posteriori para cada parâmetro do MQP e do MNL em estudo. As distribuições a priori consideradas para cada parâmetro (distribuição normal para  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\beta_1$  e  $\beta_2$ , distribuição beta para  $\beta_3$  e distribuição gama para  $\tau$ ) combinada com a função de verossimilhança, que contém a informação dos dados (distribuição normal), geraram distribuições a posteriori com formas desconhecidas para os parâmetros  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\beta_1$ ,  $\beta_2$  e  $\beta_3$  e forma conhecida para a precisão  $\tau$ . Foram utilizadas 10 mil iterações no processo MCMC para a ajuste do MQP e 200 mil iterações para o ajuste do MNL, sendo realizado o descarte de 50% das primeiras iterações nos dois modelos e saltos de 10 valores para o MQP e 200 valores para o MNL. No anexo C do presente trabalho estão apresentados os gráficos do histórico das cadeias dos processos MCMC, além dos gráficos das densidades a posteriori e gráficos de autocorrelações dos parâmetros dos modelos utilizados.

Ao verificar a evidência de convergência das cadeias do processo MCMC para os parâmetros dos modelos, apresenta-se na Tabela 8 o resultado do teste de convergência de Geweke. Tanto para os parâmetros do MQP, quanto para os parâmetros do MNL o teste apresenta Z-scores entre o intervalo de  $-1,96$  e  $1,96$ , considerando um nível de significância de 0,05, ou seja, há evidências de convergência das cadeias para todos os parâmetros.

Tabela 8 - Teste de convergência de cadeia de Geweke, modelo de regressão quadrático com platô de resposta e modelo não linear com platô de resposta, considerando homogeneidade de variância.

MQP	Parâmetros	$\alpha$	$\beta$	$\gamma$	$x_0$	P	$\tau$	$\sigma^2$
	Z-score	1,27	-0,78	0,60	0,20	-0,82	-0,48	0,58
MNL	Parâmetros	$\beta_1=P$	$\beta_2=x_0$	$\beta_3$	$\tau$	$\sigma^2$	-	-
	Z-score	1,01	0,64	-0,46	-0,62	0,61	-	-

O resumo da distribuição a posteriori dos parâmetros do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta estão apresentados na Tabela 9. As médias a posteriori dos parâmetros dos modelo MQP e MNL apresentaram um erro de Monte Carlo (EMC) baixo e menor do que o desvio padrão a posteriori de cada parâmetro. Isso indica que o número de iterações na cadeia de cada parâmetro é adequado.

O ajuste apresentou uma estimativa para o ponto de troca do MQP de  $\hat{x}_0=66,89$  ppm, que representa o nível de Zn na ração a partir do qual a deposição deste nutriente na tibia dos animais estaria estabilizada. Para o MNL, a estimativa do ponto de troca foi de  $\hat{x}_0=68,37$ . Esses valores são inferiores aos encontrados por Rezende et al. (2007) para o MQP, que foram de  $\hat{x}_0=71,80$  e semelhante ao do MNL que foi de  $\hat{x}_0=68,10$ , ajustando os mesmos modelos com platô de resposta para os teores médios de Zn na tibia em função de cada dose por meio de abordagem clássica. Isso deve-se ao fato dos autores terem utilizado a média do teor de Zn em cada uma das dosagens no ajuste e não a utilização de todas as repetições do teor em função das doses. Sendo assim, os valores extremos podem ter afetado os valores dos teores médios de Zn em cada dosagem.

Os níveis de Zn encontrados nos ajustes do presente trabalho, também foram inferiores aos encontrados por Gomes et al. (2009), que utilizaram modelos de regressão quadrático para estimar o nível adequado de Zn na ração de frangos machos e fêmeas. Os autores determinaram que a exigência ideal de Zn para ração seria de 82,20 ppm para machos e 85,70 ppm para fêmeas.

Para o platô de resposta ( $P$ ), o ajuste do modelo MQP resultou em uma estimativa do teor médio de Zn encontrado na tibia dos frangos de  $\hat{P} = 290,23$  ppm e foi superior ao encontrado por Rezende et al. (2007), que foi de 281,1 ppm utilizando o

mesmo modelo para o teor médio de Zn em cada dose do mesmo experimento. Essa seria a deposição máxima de Zn na tíbia dos frangos considerando o nível ideal de Zn na ração dos mesmos. No MNL, a estabilização na deposição do mineral foi estimada em  $\hat{P} = 290,68$ , o que foi bem próximo ao obtido por meio do ajuste do modelo MQP.

O intercepto apresentou uma média a posteriori de  $\hat{\alpha} = 208,05$  ppm, ou seja, sem adição de Zn na ração das aves, o teor de Zn na tíbia seria de aproximadamente 208,05 ppm. Cabe ressaltar que as fases analisadas correspondem à fase resposta e estável, ou seja, o experimento aconteceu após a fase inicial caracterizada por Euclides e Rostagno (2001).

Tabela 9 - Médias a posteriori, desvios padrões (DP) a posteriori e intervalo de credibilidade HPD e erro de Monte Carlo (EMC) dos parâmetros do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta, considerando homogeneidade de variância.

	Parâmetros	Média	DP	EMC	L.I	L.S
MQP	$\alpha$	208,05	5,15	0,18	197,40	217,70
	$\beta$	2,51	0,43	0,01	1,75	3,40
	$\gamma$	-0,0196	0,0062	0,0002	-0,0321	-0,0089
	$x_0$	66,89	9,72	0,29	48,99	86,77
	P	290,23	2,80	0,05	284,80	295,70
	$\tau$	0,0022	0,0003	0,000004	0,0017	0,0028
	$\sigma^2$	460,22	57,82	0,85	355,40	577,10
MNL	$\beta_1=P$	290,68	2,67	0,05	285,5	295,9
	$\beta_2=x_0$	68,37	8,25	0,18	53,49	84,77
	$\beta_3$	0,00007	0,000019	0,00007	0,00004	0,00011
	$\tau$	0,0022	0,0003	0,000002	0,0017	0,0027
	$\sigma^2$	462,18	57,01	0,32	355,3	574,6

L.I - limite inferior do intervalo HPD 95%;

L.S - limite superior do intervalo HPD 95%

Na Figura 8, tem-se o ajuste do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta para os dados do teor de Zn na tíbia dos frangos, para o caso em que foi considerado homogeneidade de variância. De acordo com o gráfico, os dois modelos apresentaram ajustes semelhantes e, com relação a dosagem

ideal na ração, o MNL apresentou intervalo de credibilidade HPD menor que no modelo MQP com um nível de credibilidade de 95%, o que indica que a probabilidade a posteriori de o parâmetro  $x_0$  estar no intervalo é de 0,95.

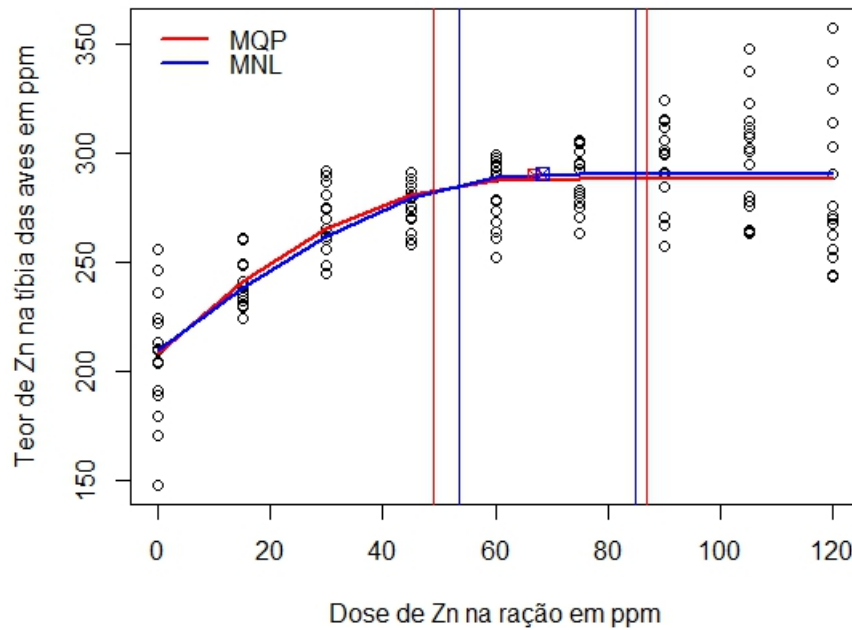


Figura 8 - Ajuste dos modelos MQP e MNL, aos dados do teor de Zinco na tíbia de frangos, no caso em que foi considerado homogeneidade de variância, e intervalos de credibilidade (linhas verticais) para o ponto de troca

#### 4.3.2 Ajuste do modelo de regressão quadrático com platô de resposta (MQP) e do modelo não linear com platô de resposta(MNL), considerando heterogeneidade de variâncias

Para o ajuste dos modelos MQP e MNL em que se considerou heterogeneidade de variâncias, procedeu-se com as mesmas distribuições a priori utilizadas na seção anterior, além dos mesmo número de iterações, descarte de 50% da cadeia e saltos 200 valores.

Na Tabela 10 é possível observar o resultado do teste de convergência das cadeias para cada parâmetro do MQP e do MNL. Segundo o teste, há evidência da convergência das cadeias do processo MCMC para todos os parâmetros estudados, pois o Z-scores



calculados são menores que 1,96 e maiores que -1,96, que são os quantis de uma normal padrão considerando um nível de significância de 0,05.

Tabela 10 - Teste de convergência de cadeia de Geweke, modelo de regressão quadrático com platô de resposta e modelo não linear com platô de resposta, considerando heterogeneidade de variâncias.

MQP	Parâmetros	$\alpha$	$\beta$	$\gamma$	$x_0$	P
	Z-score	-0,83	0,68	-0,60	-0,35	0,45
MNL	Parâmetros	$\beta_1=P$	$\beta_2=x_0$	$\beta_3$	-	-
	Z-score	0,06	0,07	0,02	-	-

O resumo a posteriori dos parâmetros do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta, composto por média, desvio padrão á posteriori e dos limites inferior e superior do intervalo de credibilidade HPD, estão apresentados na Tabela 11. As estimativas dos erros de Monte Carlo evidenciam que para os parâmetros dos modelos MQP e MNL, o número de iterações nas cadeias foram adequadas, pois as estimativas do EMC são baixas e menores que os respectivos desvios padrões a posteriori.

No ajuste do MQP, o parâmetro  $\alpha$ , que representa o intercepto, obteve uma média a posteriori de 211,88. O ponto de troca apresentou média a posteriori de  $\hat{x}_0 = 73,26\text{ppm}$  para o MQP, e para o MNL  $\hat{x}_0 = 74,03\text{ppm}$ , ou seja, estimativas bem próximas para os dois modelos ajustados e superiores as obtidas nos ajuste do MQP e MNL considerando uma variância, na seção anterior. Para estes valores obtidos de dosagem adequada de zinco na ração dos frangos, obteve-se uma deposição de zinco na tibia dos animais de  $\hat{P} = 290,12\text{ppm}$  e  $\hat{P} = 290,60\text{ppm}$ , respectivamente, para os dois modelos.

Considerar a heterogeneidade de variâncias entre os níveis do fator dose mostrou-se influenciar nas estimativas dos parâmetros dos modelos ajustados, pois houve diferença considerável nas estimativas encontradas no MQP e MNL.

Tabela 11 - Médias a posteriori, desvios padrões (DP) a posteriori, intervalo de credibilidade HPD e erro de Monte Carlo (EMC), dos parâmetros do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta, considerando uma variância para cada nível de dose.

	Parâmetros	Média	DP	EMC	L.I	L.S
MQP	$\alpha$	211,88	5,71	0,33	200,30	222,80
	$\beta$	2,19	0,39	0,02	1,46	2,96
	$\gamma$	-0,0160	0,0052	0,0003	-0,0259	-0,0071
	$x_0$	73,26	10,43	0,54	52,99	93,46
	P	290,12	3,28	0,11	284,00	296,70
MNL	$\beta_1=P$	290,56	3,06	0,07	284,60	296,50
	$\beta_2=x_0$	74,03	8,92	0,25	56,93	92,00
	$\beta_3$	0,000057	0,000016	0,00002	0,000029	0,000088

L.I - limite inferior do intervalo HPD 95%;

L.S - limite superior do intervalo HPD 95%

Os ajustes do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta aos dados do teor de Zinco na tíbia de frangos, no caso que foi considerado a heterogeneidade de variâncias entre níveis do fator dose, está representado na Figura 9. Os dois modelos apresentaram ajustes semelhantes, além disso, os ajustes de ambos parecem satisfatório. O ponto de troca, dosagem ideal na ração, apresentou menor intervalo de credibilidade HPD para o MNL com um nível de credibilidade de 95%, conforme linhas verticais no gráfico da Figura 9.

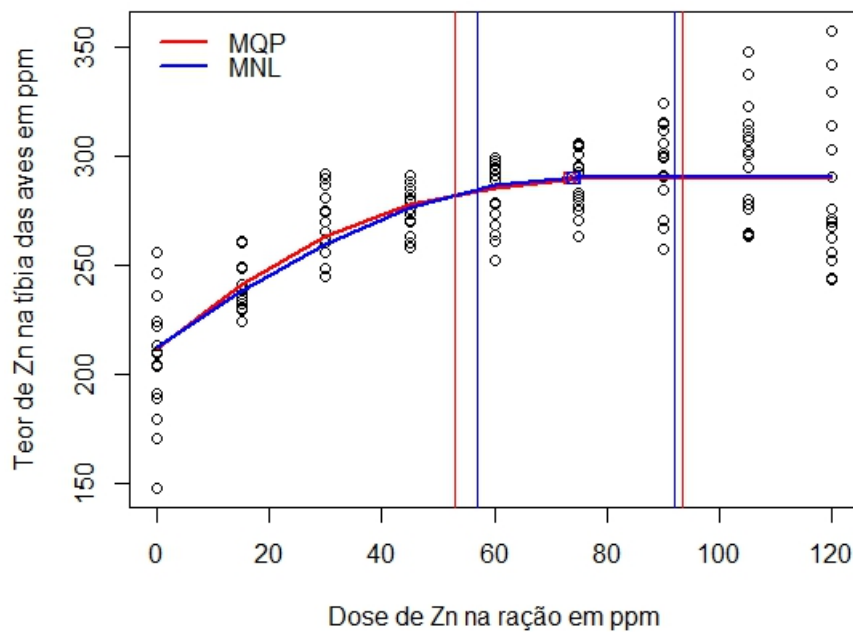


Figura 9 - Ajuste dos modelos MQP e MNL, aos dados do teor de Zinco na tíbia de frangos, no caso em que foi considerado heterogeneidade de variâncias, e intervalos de credibilidade (linhas verticais) para o ponto de troca.

#### 4.3.3 Diagnósticos dos ajustes e critério de seleção do modelo

Para avaliar e validar o ajuste dos modelos, seguem nas Figuras 10 e 11, os gráficos dos resíduos versus os valores preditos. Nos casos em que foram consideradas variâncias constantes, nota-se que há uma heterogeneidade na variação dos resíduos em cada valor predito. Já no caso em que considerou-se uma variância para o teor de Zn em cada nível de dose, percebe-se que os resíduos apresentam homogeneidade entre os valores preditos.

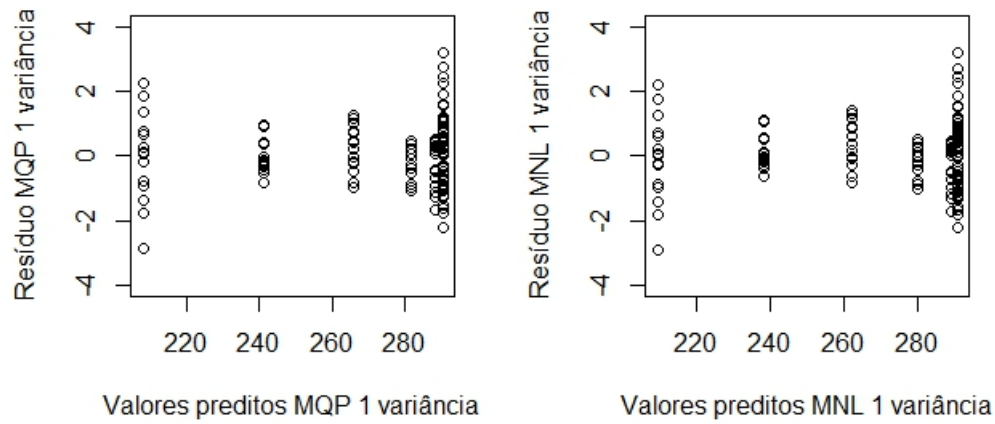


Figura 10 - Resíduos padronizados x valores preditos do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta para os casos em que considerou-se homogeneidade de variância.

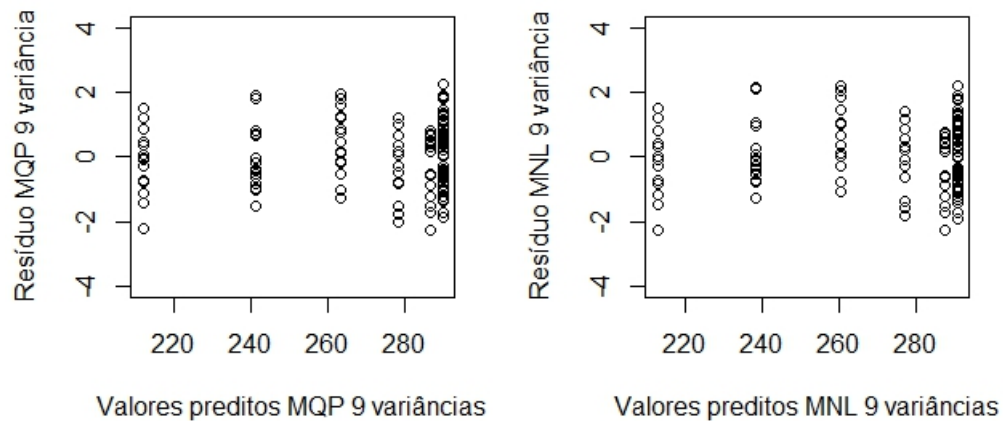


Figura 11 - Resíduos padronizados x valores preditos do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta para os casos em que considerou-se heterogeneidade de variâncias.

Comparando os ajustes dos modelos de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta por meio do critério de informação

deviance (DIC), na Tabela 12, o MQP considerando homogeneidade de variância apresentou DIC=1240. Já o MNL, também considerando homogeneidade de variância, obteve o mesmo valor de DIC=1240. No caso em que foi considerado heterogeneidade de variâncias, o MQP apresentou um DIC=1211 e o MNL um DIC=1212, ou seja, valores semelhantes. Dessa forma, pelo critério DIC, os dois modelos em que considerou-se heterogeneidade de variâncias são melhores para o ajuste aos dados da deposição de zinco na tíbia de frangos.

O resultado do DIC corrobora com os gráficos de resíduos apresentados na Figura 11, ou seja, os modelos em que foi considerado heterogeneidade de variâncias apresentaram melhor ajuste.

Tabela 12 - Critério de informação deviance (DIC) dos modelos MQP e MNL.

Modelo	Variância	DIC
MQP	1 variância	1240
MNL		1240
MQP	9 variâncias	1211
MNL		1212

## 5 CONCLUSÃO

A utilização da metodologia bayesiana, mostrou-se adequada para estimar parâmetros de modelos de regressão com platô de resposta, em diferentes tamanhos amostrais e diferentes comportamentos da variância. Com a metodologia, no estudo de simulação, foi possível determinar o ponto de troca dos segmentos ( $x_0$ ), ou seja, a quantidade adequada para dosagem do Zinco na ração dos animais, em diferentes cenários. A distribuição de  $x_0$  evidencia que quanto maior a amostra, melhor será definido seu valor médio, ou seja, a variância é menor quando se aumenta o tamanho amostral. Além disso, a deposição máxima suportada pelas aves (platô de resposta P) também foi adequada, nos diferentes cenários e ainda considerando todas as repetições em cada dosagem do experimento, e não só as médias em cada dosagem.

Sendo assim, a metodologia bayesiana, é uma alternativa aos métodos de inferência clássica na obtenção de valores ótimos, como no caso de modelos de regressão com platô de resposta.

No ajuste do modelo de regressão quadrático com platô de resposta e do modelo não linear com platô de resposta aos dados do teor de Zinco na tíbia de frangos, considerando homogeneidade e heterogeneidade de variâncias entre os níveis do fator dose, os modelos mostraram-se bem ajustados de acordo com as estatísticas e análise dos resíduos. Porém, pelo critério de qualidade de ajuste DIC, os modelos em que considerou-se heterogeneidade variâncias apresentaram menores valores da estatística DIC. Sendo assim, os modelos MQP e MNL considerando heterogeneidade de variâncias foram melhores no ajuste aos dados e são mais adequados para o estudo.



## REFERÊNCIAS

- BACON, D. W.; WATTS, D. G. Estimating the transition between two intersecting straight lines. **Biometrika**, v. 58, n. 3, p. 525–534, 1971.
- BARTEN, A. P.; BRONSARD, L. S. Two-stage least-squares estimation with shifts in the structural form. **Econometrica**, v. 38, p. 938–941, 1970.
- BERTECHINI, A. G.; HOSSAIN, S. M.; NOBRE, P. T. Efeito do desempenho níveis de zinco no desempenho e composição mineral do tecido em frangos de corte. In: **Anais: Reunião Anual da Sociedade Brasileira de Zootecnia**. [S.l.: s.n.], 1992.
- BOX, G. E. P.; TIAO, G. C. **Bayesian Inference in Statistical Analysis**. [S.l.]: John Wiley & Sons, 1992. 588 p.
- BROCHI, A. L. C; DINIZ, C. A. R. Desempenho de métodos de estimação em modelos de regressão linear simples segmentada. **Revista de Matemática e Estatística**, v. 19, p. 325–347, 2001.
- CARLIN, B. P.; GELFAND, A. E.; SMITH, A. F. M. Hierarchical bayesian analysis of changepoint problems. **Royal Statistical Society**, v. 41, n. 2, p. 389–405, 1992.
- CASELLA, G.; GEORGE, E. I. Explaining the gibbs sampler. **The American Statistician**, v. 46, n. 3, p. 167–174, 1992.
- CHEN, C. W. S.; CHAN, J. S. K.; GERLACH, R.; HSIEH, W. Y. L. A comparison of estimators for regression models with change points. **Statistics and Computing**, v. 21, p. 395–414, 2011.
- CHEN, M.; BAN-WEISS, G. A.; SANDERS, K. T. The role of household level electricity data in improving estimates of the impacts of climate on building electricity use. **Energy and Buildings**, 2018.
- CHINCHOY, J. H.; BROEMELING, L. D. Some bayesian inference for a changing linear model. **Technometrics**, v. 22, n. 1, p. 71 – 78, 1980.
- COELHO, L. S. S.; COSTA, P. M. A.; SILVA, M. A.; PEREIRA, J. A. A.; ROSTAGNO, H. S.; BARBOSA, H. P. Modelos para estimar exigências nutricionais em suínos. **Revista Brasileira de Zootecnia**, v. 16, n. 1, p. 102–110, 1987.
- COLDEBELLA, A.; AVILA, V. S. de; BRUM, P. A. R. de; JUNIOR, W. B.; PAULA, A. Modelos de regressão para determinar a estabilização da variabilidade de dados de energia metabolizável em aves. In: **Simpósio de Estatística Aplicada à Experimentação Agrônômica (SEAGRO) 11. Reunião da Região Brasileira da Sociedade Internacional de Biometria 50**. [S.l.: s.n.], 2005. p. 183.
- DINIZ, C. A. R.; MILAN, L. A.; MAZUCHELI, J. Bayesian inference of linear segmented regression model. **Brazilian Journal of Probability and Statistics**, v. 17, p. 1–16, 2003.
- EHLERS, R. S. **Introdução à bayesiana**. 2007. Disponível em: <<http://www.leg.ufpr.br/paulojus/CE227/ce227.pdf>>.



EUCLYDES, R. F.; ROSTAGNO, H. S. Estimativa dos níveis nutricionais via experimentos de desempenho. In: **Workshop Latino Americano Anjinomoto Biolatina**. [S.l.: s.n.], 2001.

FERREIRA, P. E. A bayesian analysis of a switching regression model: Known number of regimes. **Journal of the American Statistical Association**, v. 70, n. 350, p. 370–374, 1975.

GAMERMAN, D.; LOPES, H. F. **Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference**. 2. ed. [S.l.]: Chapman & Hall book, 2006.

GELFAND, A. E. Gibbs sampling. **Journal of the American Statistical Association**, v. 95, n. 452, p. 1300–1304, 2000.

GELFAND, A. E.; HILLS, S. E.; RACINE-POON, A.; SMITH, A. F. M. Illustration of bayesian inference in normal data models using gibbs sampling. **Journal of the American Statistical Association**, v. 85, n. 412, p. 972–985, 1990.

GELFAND, A. E.; SMITH. Sampling based approaches for calculating marginal densities. **Journal of the American Statistical Association**, v. 85, n. 410, p. 398–409, 1990.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; DUNSON, D. B.; VEHTARI, A.; RUBIN, D. B. **Bayesian data analysis**. third. [S.l.]: Chapman & Hall book, 2014.

GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. **Statistical Science**, v. 7, n. 4, p. 457–511, 1992.

GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 6, p. 721–741, 1984.

GEWEKE, J. Evaluating the accuracy of sampling based approaches to calculating posterior moments. **Bayesian Statistics**, v. 4, p. 625–631, 1992.

GOLDFELD, S. M.; QUANDT, R. E. The estimation of structural shifts by switching regressions. **Annals of Economic and Social Measurement**, v. 2, p. 475–485, 1973.

GOMES, P. C.; RIQUEIRA, D. C. M.; BRUMANO, G.; ALBINO, L. F. T.; ROSTAGNO, H. S.; SCHMIDT, M. Níveis nutricionais de zinco para frangos de corte machos e fêmeas nas fases de crescimento e terminação. **Revista Brasileira de Zootecnia**, v. 38, n. 9, p. 1719–1725, 2009.

GOMES, P. C.; RIQUEIRA, D. C. M.; ROSTAGNO, H. S.; ALBINO, L. F. T.; B., G.; SCHMIDT, M. Exigência nutricionais de zinco para frangos de corte machos e fêmeas na fase inicial. **Revista Brasileira de Zootecnia**, v. 37, n. 1, p. 79–83, 2008.

GUARCONI, R. C.; SOUZA, J. L. de; FAVARATO, L. F.; ANGELETTI, M. da P.; BAHENSE, D. V. Determinação do tamanho ótimo de parcela experimental para experimentos com repolho utilizando simulação e métodos de estimação. **Revista Científica Intelletto**, v. 2, n. 2, p. 79–87, 2017.

HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. **Biometrika**, v. 57, n. 1, p. 97–109, 1970.

HEIDELBERG, P.; WELCH, P. Simulation run lenght control in the presence of an initial transient. **Operations**, v. 31, n. 6, p. 1109–1144, 1983.

HOLBERT, D. A bayesian analysis of a switching linear model. **Journal of Econometrics**, v. 19, p. 77–87, 1982.

JULIOUS, S. A. Inference and estimation in a chandepoint regression problem. **The Statistician**, v. 50, n. Parte 1, p. 51–61, 2001.

KAPS, M.; LAMBERSON, W. R. **Biostatistics for animal science**. [S.l.]: CABI, 2004.

KIM, E.; REW, H. J.; SHIN, T. K.; CHO, H. M.; WICKRAMASURIYA, S. S.; YI, Y. J.; JEONG, J.; CHOI, I.; HEO, J. M. Standard body weight and serum estradiol and progesterone concentrations in response to total lysine content in female broiler breeders from 14 to 42 days after hatch. **Brazilian Journal of Poultry Science**, v. 20, n. 1, p. 71–78, 2018.

KÉRY, M. **Introduction to WinBugs for ecologists: A Bayesian approach to regression, ANOVA, mixed models and related analyses**. [S.l.]: Library of Congress Cataloging in Publication Data, 2010.

LERMAN, P. M. Fitting segmented regression models by grid search. **Journal of the Royal Statistical Society. Series C (applied Statistics)**, v. 29, n. 1, p. 77–84, 1980.

LERMAN, P. M.; BIE, S. W. Problems in determining the best levels of essential nutrients in feedingstuffs. **The Journal of Agricultural Science**, v. 84, p. 459–468, 1975.

METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H. Equation of state calculations by fast computing machines. **The Journal of Chemical Physics**, v. 21, n. 6, p. 1087–1092, 1953.

MOREIRA, J. M.; MELO, A. F.; OLIVEIRA, J. M. de; ATAIDES, D. S.; RIBEIRO, M. C.; BORTOLINI, J. Parcela ótima para a cultura do cafeeiro obtido por simulação de dados com variâncias conhecidas. **Publicações em Medicina Veterinária e Zootecnia**, v. 10, n. 9, 2016.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. São Paulo: Edgard Blucher, 2004. 535 p.

MUGGEO, V. M. R. Estimating regression models with unknown break-points. **Statistics in Medicine**, v. 22, p. 3055–3071, 2003.

NOGUEIRA, D. A.; SáFADI, T.; FERREIRA, D. F. Avaliação de critérios de convergência para o método de monte carlo via cadeias de markov. **Revista Brasileira de Estatística**, v. 65, n. 224, p. 59–88, 2004.

OLIVEIRA, A. de M.; FURLAN, A. C.; MURAKAMI, A. E.; MOREIRA, I.; SCAPINELLO, C.; MARTINS, E. N. Exigência nutricional de lisina para codornas japonesas (*coturnix coturnix japonica*) em postura. **Revista Brasileira de Zootecnia**, v. 28, n. 5, p. 1050–1053, 1999.

PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. [S.l.]: Fundação Calouste Gulbenkian, 2003.

PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B.; SILVA, G. **Estatística bayesiana**. segunda. [S.l.]: Fundação, 2018.

PEIXOTO, A. P. B.; FARIA, G. A.; MORAIS, A. R. de. Modelos de regressão com platô na estimativa do tamanho de parcelas em experimento de conservação in vitro de maracujazeiro. **Ciência Rural**, v. 41, n. 11, p. 1907–1913, 2011.

PESTI, G. M.; VEDENOV, D.; CASON, J. A.; BILLARD, L. A comparison of methods to estimate nutritional requirements from experimental data. **British Poultry Science**, v. 50, n. 1, p. 16–32, 2009.

PLUMMER, M.; BEST, N.; COWLES, K.; VINES, K. Coda: Convergence diagnosis and output analysis for mcmc. **R News**, v. 6, n. 1, p. 7–11, 2006. Disponível em: <<https://journal.r-project.org/archive/>>.

PORTZ, L.; DIAS, C. T. dos S.; CYRINO, J. E. P. Regressão segmentada como modelo na determinação de exigências nutricionais de peixes. **Scientia Agricola**, v. 57, n. 4, p. 601–607, 2000.

QUANDT, R. E. The estimation of the parameters of a linear regression system obeying two separate regimes. **Journal of the American Statistical Association**, v. 53, n. 284, p. 873–880, 1958.

QUANDT, R. E. Tests of the hypothesis that a linear regression system obeys two separate regimes. **Journal of the American Statistical Association**, v. 55, n. 290, p. 324–330, 1960.

QUANDT, R. E. A new approach to estimating switching regressions. **Journal of the american statistical association**, v. 67, n. 338, p. 306–310, 1972.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2016. Disponível em: <<https://www.R-project.org/>>.

RAFTERY, A. L.; LEWIS, S. Comment: one long run with diagnostics: implementation strategies for markov chain monte carlo. **Statistical Science**, v. 7, n. 4, p. 493–497, 1992.

REZENDE, D. M. L. C. **Ajuste de modelos de platô de resposta aplicados ao estudo de exigência nutricional de frangos de corte**. 2002. Dissertação (Mestrado) — UFLA, 2002.

REZENDE, D. M. L. C.; MUNIZ, J. A.; FERREIRA, D. F.; SILVA, F. F. e; AQUINO, L. H. de. Ajuste de modelos de platô de resposta para a exigência de zinco em frangos de corte. **Ciência e Agrotecnologia**, v. 31, n. 2, p. 468–478, 2007.

ROBBINS, K. R. A method, sas program, and example for fitting the broken-line to growth data. **Research reports: University of Tennessee Agricultural experiment station**, v. 86, n. 09, p. 8, 1986.

ROBERTS, G.O.; SMITH, A. F. M. Simple conditions for the convergence of the gibbs sampler and metropolis-hastings algorithms. **Stochastic Processes and their applications**, v. 49, p. 207–216, 1994.

ROSTAGNO, H. S.; BUNZEN, S.; SAKOMURA, N. K.; ALBINO, L. F. T. Avanços metodológicos na avaliação de alimentos e de exigências nutricionais para aves e suínos. **Revista Brasileira de Zootecnia**, v. 36, n. suplemento especial, p. 295–304, 2007.

RUNHO, R. C.; GOMES, P. C.; ROSTAGNO, H. S.; ALBINO, L. F. T.; LOPES, P. S.; POZZA, P. C. Exigência de fósforo disponível para frangos de corte machos e fêmeas de 1 a 21 dias de idade. **Revista Brasileira de Zootecnia**, v. 30, n. 1, p. 187–196, 2001.

SCHABENBERGER, O.; PIERCE, F. J. **Contemporary Statistical models for the plant and soil sciences**. [S.l.]: CRC Press, 2002.

SCHMITTDIEL, J. A.; ADAMS, S. R.; GOLER, N.; SANNA, R. S.; BOCCIO, M.; BEL-LAMY, D. J.; BROWN, S. D.; NEUGEBAUER, R. S.; FERRARA, A. The impact of telephonic wellness coaching on weight loss: A natural experiments for translation in diabetes (next-d) study. **Journal of the Obesity Society**, v. 25, n. 2, p. 352–356, 2017.

SEBER, G. A. F.; WILD, C. J. **Nonlinear Regression**. Auckland, New Zealand: John Wiley & Sons, 1989. 768 p.

SHABAN, S. A. Change point problem and two phase regression: an annotated bibliography. **International Statistical Review**, v. 48, n. 1, p. 83–93, 1980.

SILVA, J. G. C. da. Análise da adaptabilidade através de regressão linear segmentada: Aplicação. **Pesquisa Agropecuária Brasileira**, v. 30, n. 4, p. 449–462, 1995.

SILVA, J. G. C. da. Análise da adaptabilidade através de regressão linear segmentada: Fundamentos. **Pesquisa Agropecuária Brasileira**, v. 30, n. 4, p. 435–448, Abr 1995.

SILVA, J. G. C. da. Análise da adaptabilidade por regressão segmentada com estimação da junção dos segmentos. **Pesquisa Agropecuária Brasileira**, v. 33, n. 7, p. 1013–1029, Jul 1998.

SIQUEIRA, J. C. de; SAKOMURA, N. K.; NASCIMENTO, D. C. N.; FERNANDES, J. B. K. Modelos matemáticos para estimar as exigências de lisina digestível para aves de corte isa label. **Revista Brasileira de Zootecnia**, v. 38, n. 9, p. 1732–1737, 2009.

SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; LINDE, A. van der. Bayesian measures of model complexity and fit. **Royal Statistical Society**, v. 64, n. part 4, p. 583–639, 2002.

STURTZ, S.; LIGGES, U.; GELMAN, A. R2winbugs: A package for running winbugs from r. **Journal of Statistical Software**, v. 12, n. 3, p. 1–16, 2005. Disponível em: <<http://www.jstatsoft.org>>.

## ANEXOS

**Anexo A - Rotina no R: Estudo de simulação**

```
#####
##### ROTINA NO R – Estudo de simulacao #####
#####
rm(list=ls(all=TRUE))

#### Modelo quadratico com plato de resposta – cenario 1

# guardando os resultados na matriz
rep_model = 100 # quantidade de alphas, betas, ...
M_result = matrix(0,rep_model,6) #linhas=rep_model=100 estimativas
#para cada parametro e 6 colunas, uma para cada parametro.

# amostras de y nas colunas
MY = matrix(0,144,rep_model)
# numero de linhas = 144 = 9 doses x 16 aves/ rep_model colunas

# valores reais
alpha=208.44265
beta=2.02291
gama=-0.01409
x0=-0.5*(beta/gama)
# doses
dexp1 <- read.delim("~/Arquivos/ESALQ/doutorado/plato_R/inicial/
dexp1.txt")

#### Bayesiana
model <- function() {
#      definicao das distribuicoes a priori
tau ~ dgamma(0.001,0.001)
alpha ~ dnorm(0.0,1.0E-6)
```

```

beta ~ dnorm(0.0,1.0E-6)
gama ~ dnorm(0.0,1.0E-6)
x0 <- -beta/(2*gama)
sigma2 <- 1/tau
for( i in 1 : N ) {
  for(j in 1:K) {
    y[i,j] ~ dnorm(mu[i,j],tau)
    mu[i,j] <- (alpha + beta*x[j] + gama*pow(x[j],2))*step(x0-x[j])+
      (alpha + beta*x0 + gama*pow(x0,2))*step(x[j]-x0)
  }
}

```

```

library(R2OpenBUGS)
model.file <- file.path(tempdir(), "model.txt")
model.file <- file.path("C:/Users/Rick/Documents/Arquivos/ESALQ/
doutorado/Bayesiana","model.txt")
write.model(model, model.file)

```

```

for(R in 1:rep_model){
  erro=rnorm(144,0,8)
  dose=dexp1$dose
  dose=sort(dose) # colocando as dosagens em ordem crescente

```

```

y=NULL
y[1:64]=alpha+beta*dose[1:64]+gama*dose[1:64]^2+erro[1:64]
y[65:144]=alpha+beta*x0+gama*x0^2+erro[65:144]
MY[,R] = y

```

```

YY=matrix(y,16,9,byrow = F)

```

```

data=list(N = 16, K=9, x = c(0,15,30,45,60,75,90,105,120),

```

```
y = YY)
```

```
params <- c("alpha","beta", "gama", "x0", "tau", "sigma2")
inits<-function(){list(alpha=207.4,beta=1.93,gama=-0.0129,tau=0.002)}
out <- bugs(data, inits, params,model.file, n.iter=10000,
n.chains = 1, n.thin = 10)
```

```
Resultado=data.frame(out$mean["alpha"],
out$mean["beta"],
out$mean["gama"],
out$mean["x0"],
out$mean["tau"],
out$mean["sigma2"])
```

```
M_result[R,] = as.numeric(Resultado)
}
```

```
### Modelo quadratico com plato de resposta - cenario 4
rm(list=ls(all=TRUE))
```

```
rep_model = 100 #
M_result169 = matrix(0,rep_model,23)
```

```
# amostras de y nas colunas
M_Y169 = matrix(0,144,rep_model) # numero de linhas = 144 e 9
#doses x 16 aves/ rep_model colunas
```

```
# valores reais
alpha=208.44265
beta=2.02291
gama=-0.01409
```

```
x0=-0.5*(beta/gama)
```

```
#### Bayesiana
```

```
model <- function() {
```

```
  for(j in 1:9){
```

```
    tau[j] ~ dgamma(0.001,0.001)
```

```
    sigma2[j] <- 1/tau[j]
```

```
  }
```

```
  alpha ~ dnorm(0.0,1.0E-6)
```

```
  beta ~ dnorm(0.0,1.0E-6)
```

```
  gama ~ dnorm(0.0,1.0E-6)
```

```
  x0 <- -beta/(2*gama)
```

```
  for( i in 1 : N ) {
```

```
    for(j in 1:K) {
```

```
      y[i,j] ~ dnorm(mu[i,j],tau[j])
```

```
      mu[i,j] <- (alpha + beta*x[j] + gama*pow(x[j],2))*step(x0-x[j])+
```

```
      (alpha + beta*x0 + gama*pow(x0,2))*step(x[j]-x0)
```

```
    }
```

```
  }
```

```
}
```

```
library(R2OpenBUGS)
```

```
model.file <- file.path(tempdir(), "model.txt")
```

```
model.file <-file.path("C:/Users/Rick/Documents/Arquivos/ESALQ/
doutorado/Bayesiana","model.txt")
```

```
write.model(model, model.file)
```

```
for(R in 1:rep_model){
```

```
  y=NULL
```

```
  dose0=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)
```

```
  dose15=c(15,15,15,15,15,15,15,15,15,15,15,15,15,15,15,15)
```



```
dose30=c(30,30,30,30,30,30,30,30,30,30,30,30,30,30,30,30)
dose45=c(45,45,45,45,45,45,45,45,45,45,45,45,45,45,45,45)
dose60=c(60,60,60,60,60,60,60,60,60,60,60,60,60,60,60,60)
dose75=c(75,75,75,75,75,75,75,75,75,75,75,75,75,75,75,75)
dose90=c(90,90,90,90,90,90,90,90,90,90,90,90,90,90,90,90)
dose105=c(105,105,105,105,105,105,105,105,105,105,105,105,105,105,105,105,105,105,105)
dose120=c(120,120,120,120,120,120,120,120,120,120,120,120,120,120,120,120,120,120,120)
```

```
erro1=rnorm(16,0,10.96)
erro2=rnorm(16,0,8.18)
erro3=rnorm(16,0,6.88)
erro4=rnorm(16,0,4.09)
erro5=rnorm(16,0,6.83)
erro6=rnorm(16,0,7.83)
erro7=rnorm(16,0,8.83)
erro8=rnorm(16,0,7.83)
erro9=rnorm(16,0,9.83)
```

```
y0=alpha+beta*dose0+gama*dose0^2+erro1
y15=alpha+beta*dose15+gama*dose15^2+erro2
y30=alpha+beta*dose30+gama*dose30^2+erro3
y45=alpha+beta*dose45+gama*dose45^2+erro4
y60=alpha+beta*x0+gama*x0^2+erro5
y75=alpha+beta*x0+gama*x0^2+erro6
y90=alpha+beta*x0+gama*x0^2+erro7
y105=alpha+beta*x0+gama*x0^2+erro8
y120=alpha+beta*x0+gama*x0^2+erro9
```

```
y=c(y0,y15,y30,y45,y60,y75,y90,y105,y120)
```

```
M_Y169[,R] = y
```

```
YY=matrix(y,16,9,byrow = F)
```

```
data=list(N = 16, K=9, x = c(0,15,30,45,60,75,90,105,120),
y = YY)
```

```
params <- c("alpha","beta", "gama", "x0", "tau","sigma2")
inits <- function() { list(alpha=207.4,beta=1.93,gama=-0.0129,
tau=c(0.0005127146, 0.0096843088, 0.0014566828, 0.0011327688,
0.0003911969, 0.0006729423, 0.0039, 0.0017, 0.001653)) }
out <- bugs(data, inits, params,model.file, n.iter=10000,
n.chains = 1, n.thin = 10)
```

```
M_result169[R, ] = unlist(out$mean)
}
```

## Anexo B - Gráficos e tabelas extras da análise de simulação

Tabela 13 - Médias das estimativas dos parâmetros de precisão do modelo de regressão quadrático com platô de resposta e vieses médios nos ajustes dos cenários 4, 5 e 6.

Parâmetro	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$	$\tau_6$	$\tau_7$	$\tau_8$	$\tau_9$
Valor paramétrico	0,0083	0,0149	0,0211	0,0598	0,0214	0,0163	0,0128	0,0163	0,0104
n=16	0,0097 (0,0014)	0,0167 (0,0018)	0,0253 (0,0041)	0,0682 (0,0084)	0,0233 (0,0019)	0,0197 (0,0034)	0,0152 (0,0024)	0,0171 (0,0008)	0,0115 (0,0012)
n=30	0,0086 (0,0003)	0,016 (0,0012)	0,0226 (0,0015)	0,0638 (0,0040)	0,0212 (-0,0002)	0,0183 (0,0020)	0,0138 (0,0010)	0,0169 (0,0006)	0,0108 (0,0005)
n=100	0,0086 (0,0003)	0,0155 (0,0005)	0,0212 (0,00009)	0,0610 (0,0013)	0,0206 (-0,0008)	0,0171 (0,0008)	0,0132 (0,0004)	0,0167 (0,0004)	0,0106 (0,0002)

Tabela 14 - Médias das estimativas dos parâmetros de precisão do modelo não linear com platô de resposta e vieses médios dos ajustes para os cenários 4, 5 e 6.

Parâmetro	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$	$\tau_6$	$\tau_7$	$\tau_8$	$\tau_9$
Valor paramétrico	0,0083	0,0149	0,0211	0,0598	0,0214	0,0163	0,0128	0,0163	0,0104
n=16	0,0096 (0,0013)	0,0175 (0,0025)	0,0260 (0,0049)	0,0703 (0,0106)	0,0262 (0,0047)	0,0182 (0,0019)	0,0140 (0,0012)	0,0188 (0,0025)	0,0119 (0,0015)
n=30	0,0088 (0,0005)	0,0158 (0,0009)	0,0230 (0,0019)	0,0649 (0,0051)	0,0227 (0,0013)	0,0164 (0,00008)	0,0136 (0,0007)	0,0170 (0,0007)	0,0109 (0,0006)
n=100	0,0084 (0,00008)	0,0150 (0,00005)	0,0213 (0,0002)	0,0607 (0,0009)	0,0215 (0,00005)	0,0169 (0,0006)	0,0130 (0,0001)	0,0166 (0,0003)	0,0106 (0,0002)

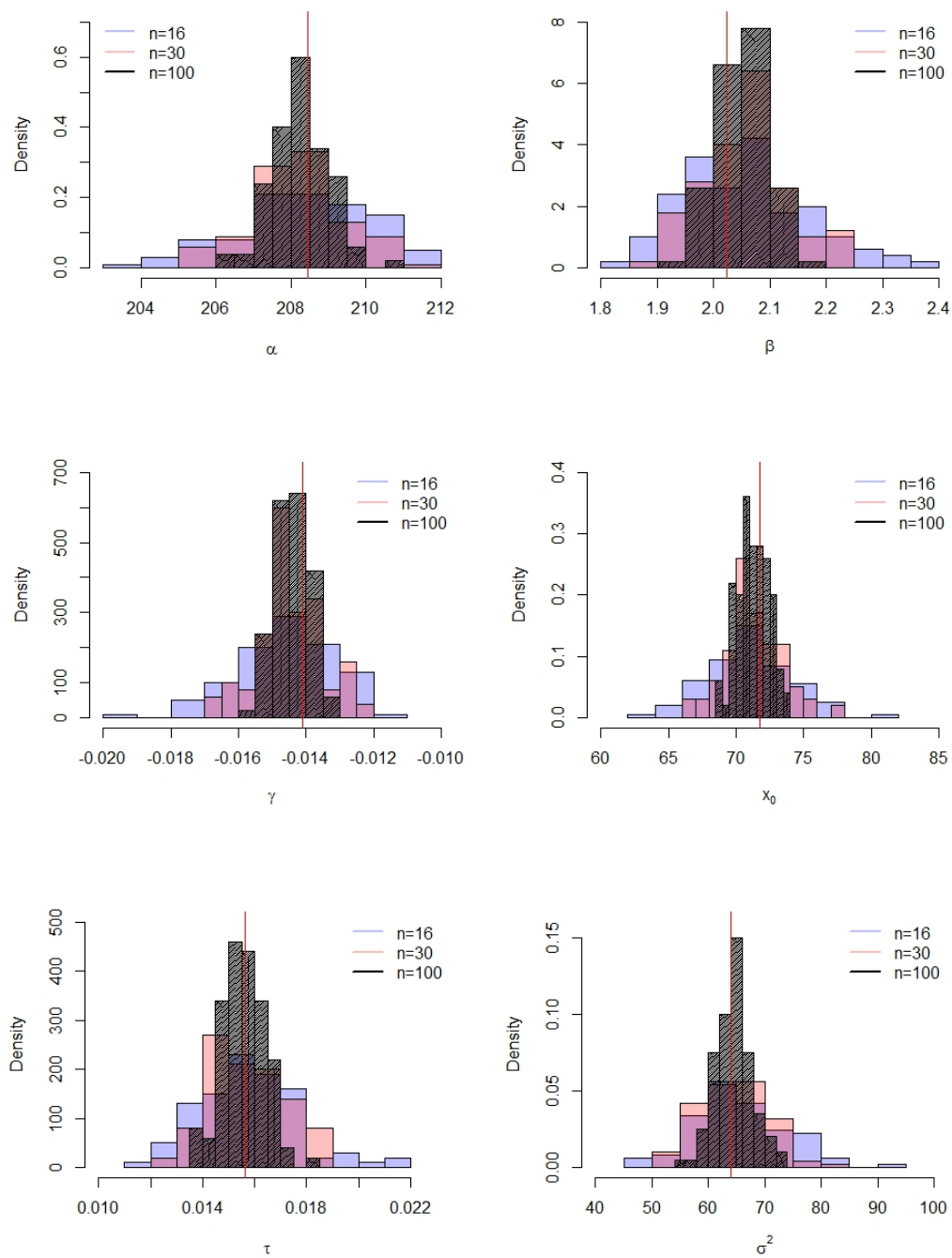


Figura 12 - Histograma das estimativas dos parâmetros do modelo de regressão quadrático com platô de resposta para os cenários 1, 2 e 3 simulados.

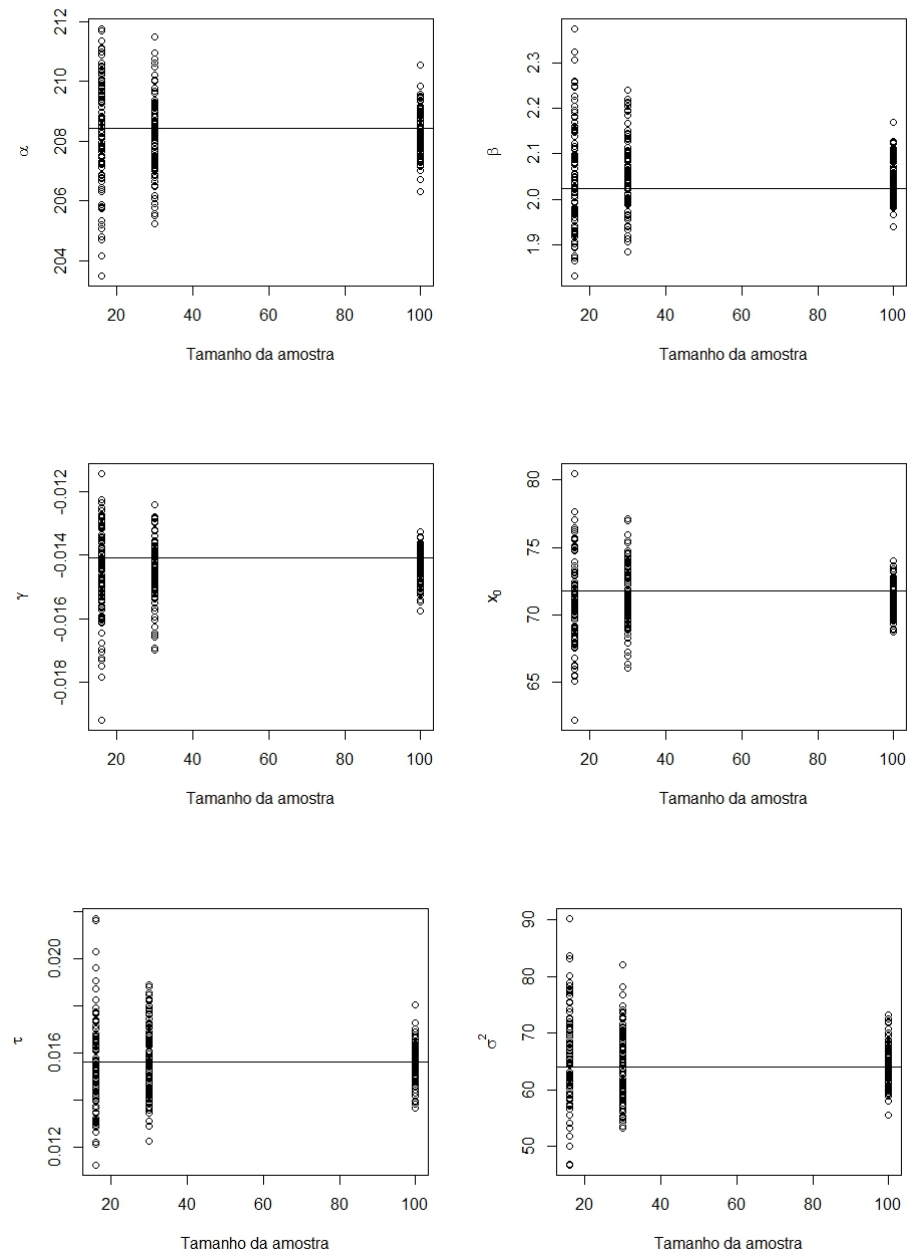


Figura 13 - Estimativas dos parâmetros do modelo de regressão quadrático com platô de resposta, em função do tamanho amostral na simulação.

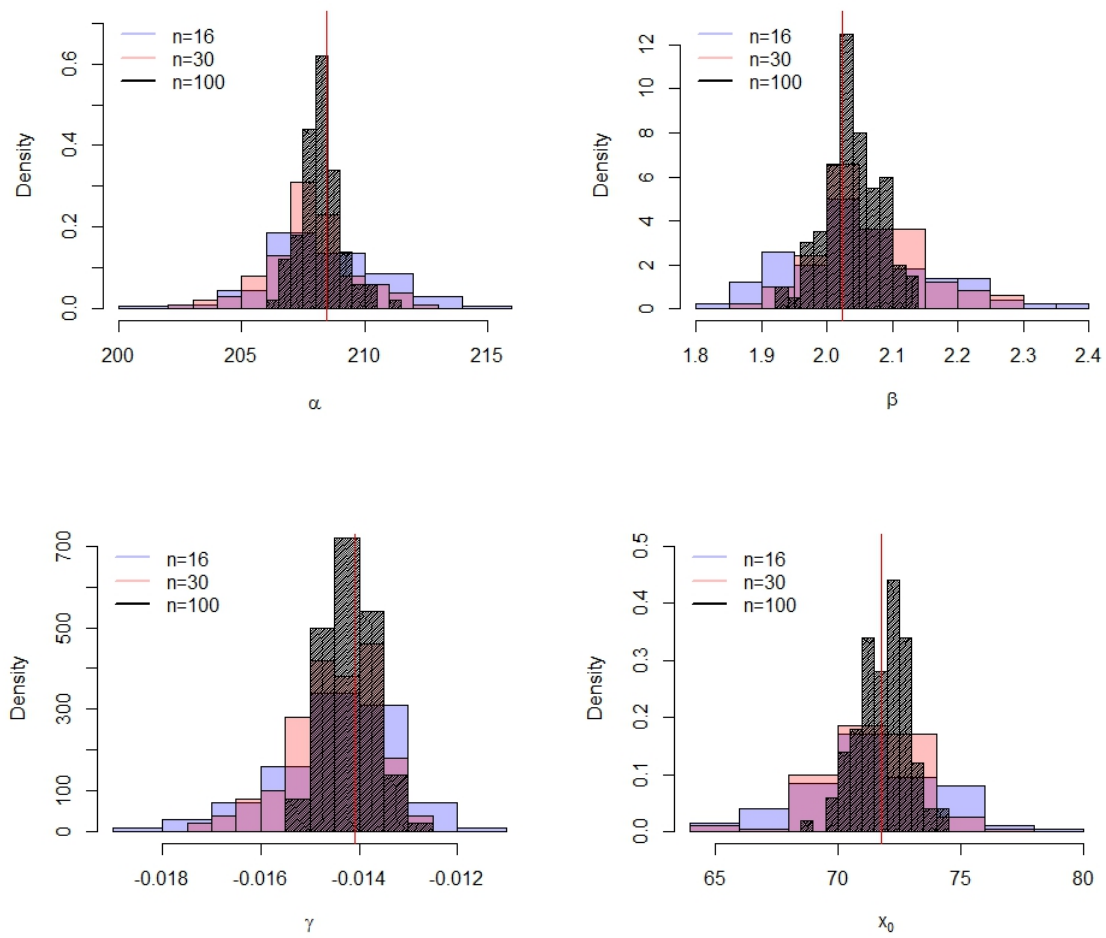


Figura 14 - Histograma das estimativas dos parâmetros do modelo de regressão quadrático com platô de resposta para os cenários 4, 5 e 6 simulados.

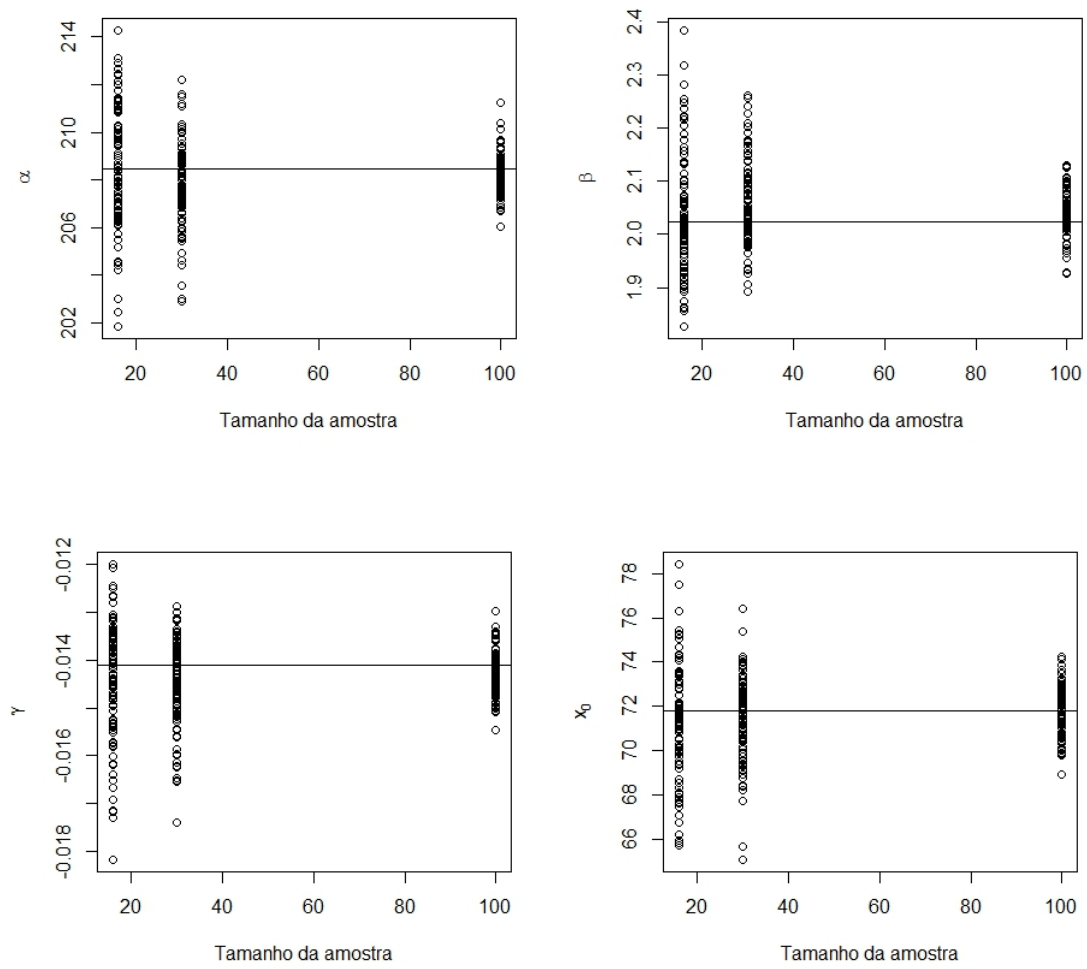


Figura 15 - Estimativas dos parâmetros do modelo de regressão quadrático com platô de resposta, em função do tamanho amostral na simulação.



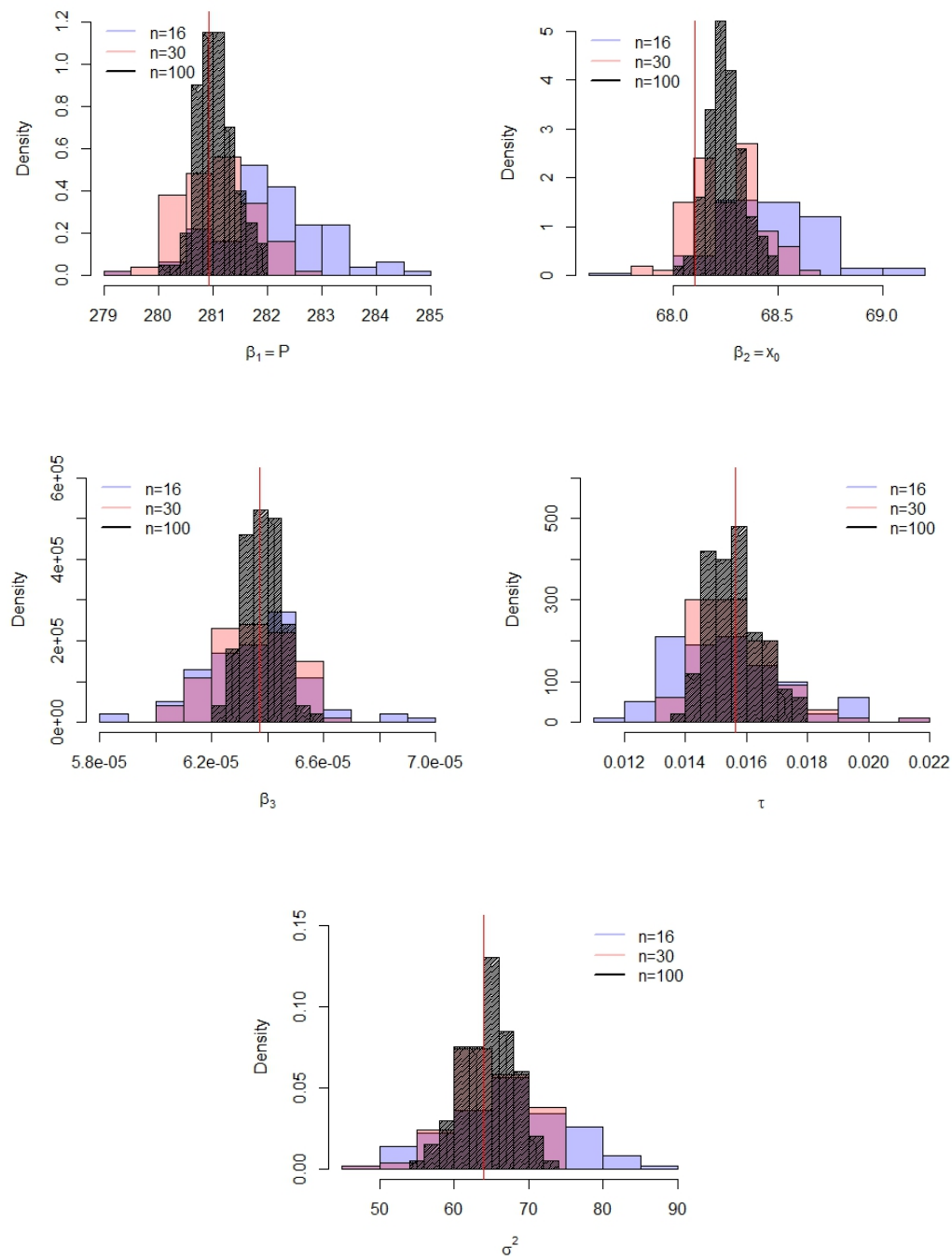


Figura 16 - Histograma das estimativas dos parâmetros do modelo não linear com platô de resposta para os cenários 1, 2 e 3 simulados.

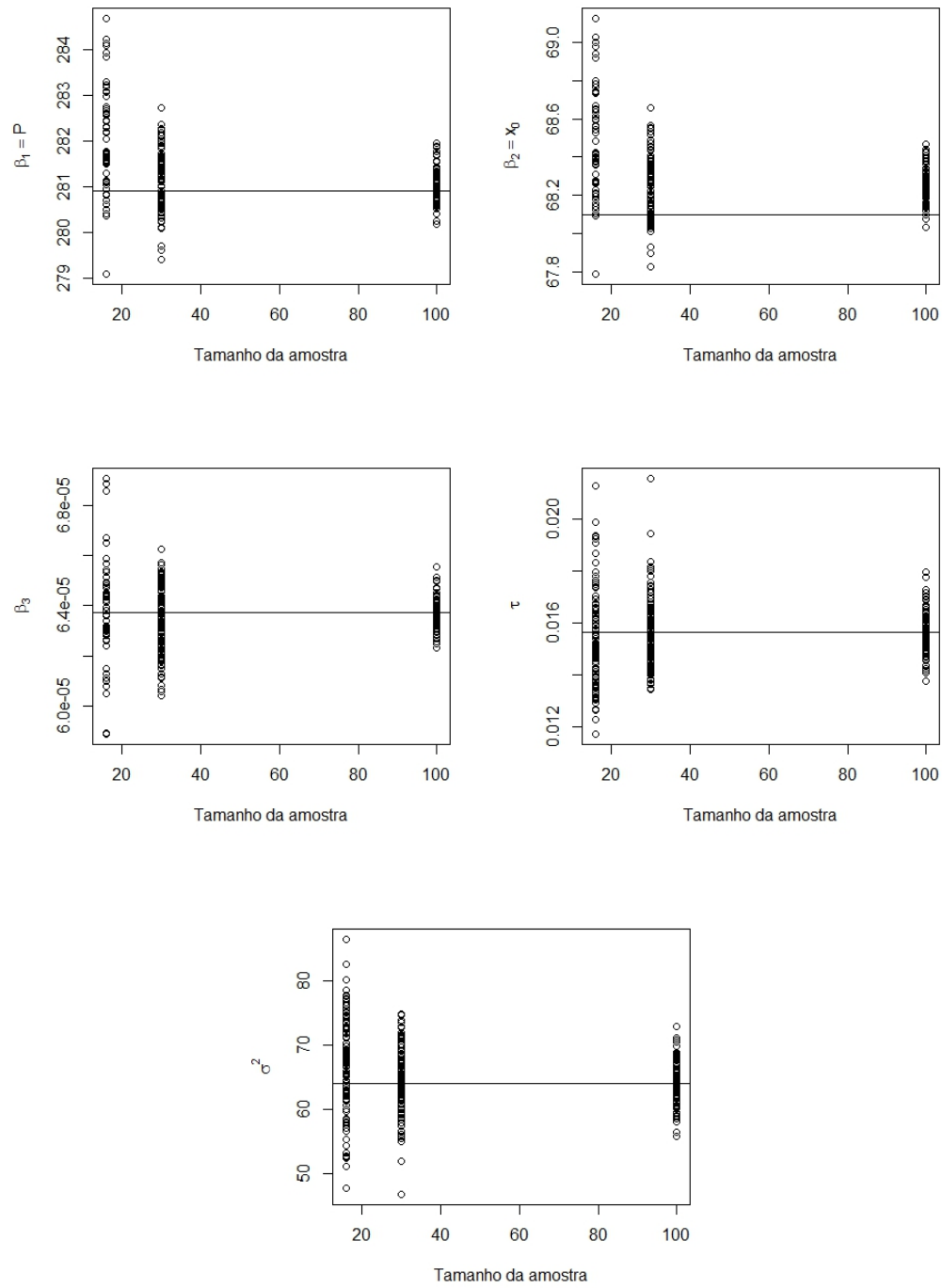


Figura 17 - Estimativas dos parâmetros do modelo não linear com platô de resposta, em função do tamanho amostral na simulação.

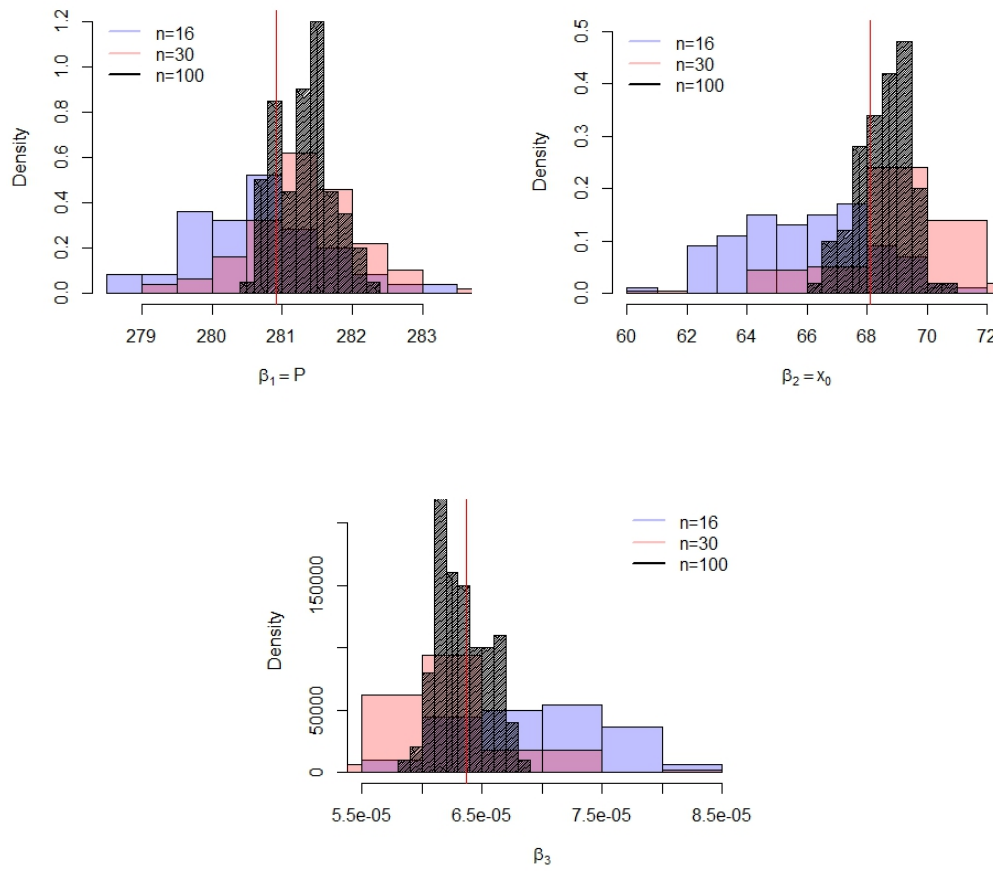


Figura 18 - Histograma das estimativas dos parâmetros do modelo não linear com platô de resposta para os cenários 4, 5 e 6 simulados.

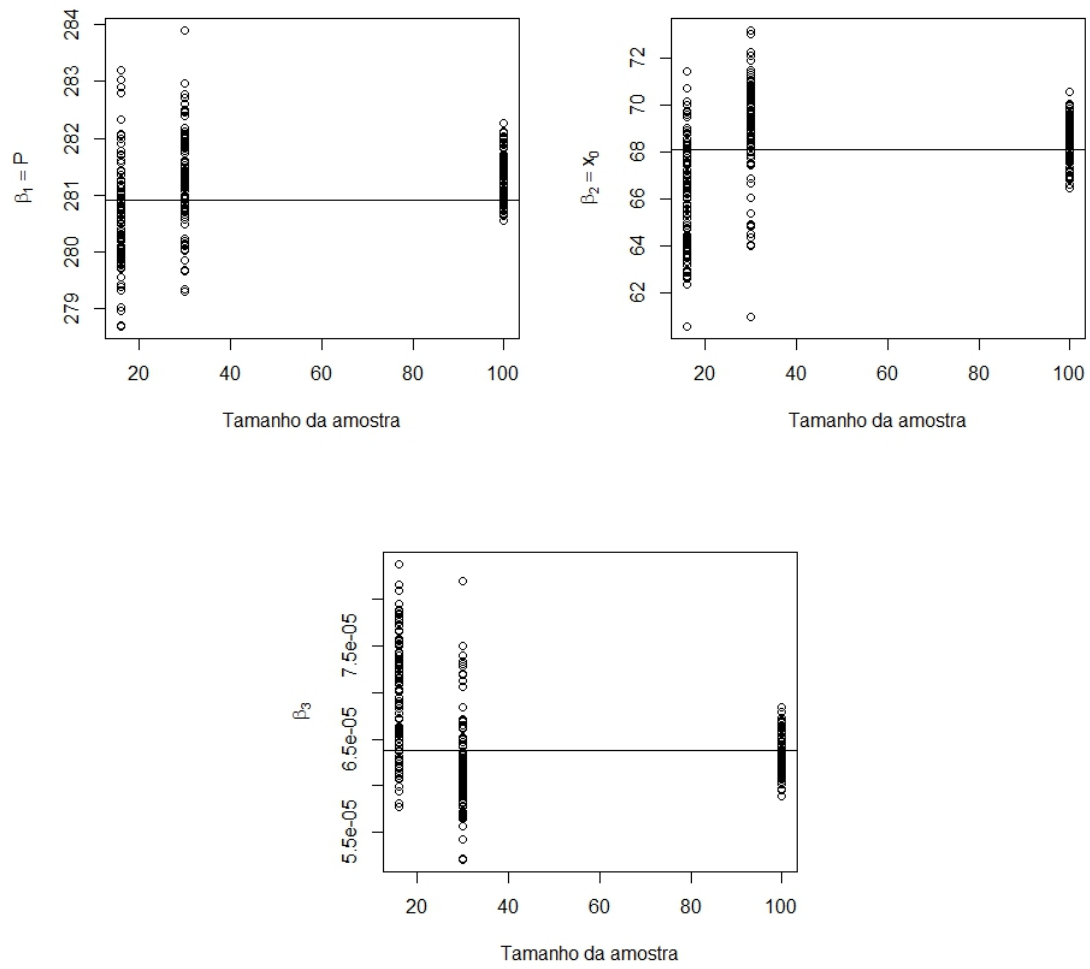


Figura 19 - Estimativas dos parâmetros do modelo não linear com platô de resposta, em função do tamanho amostral na simulação.

## Anexo C - Gráficos auxiliares da análise dos dados reais

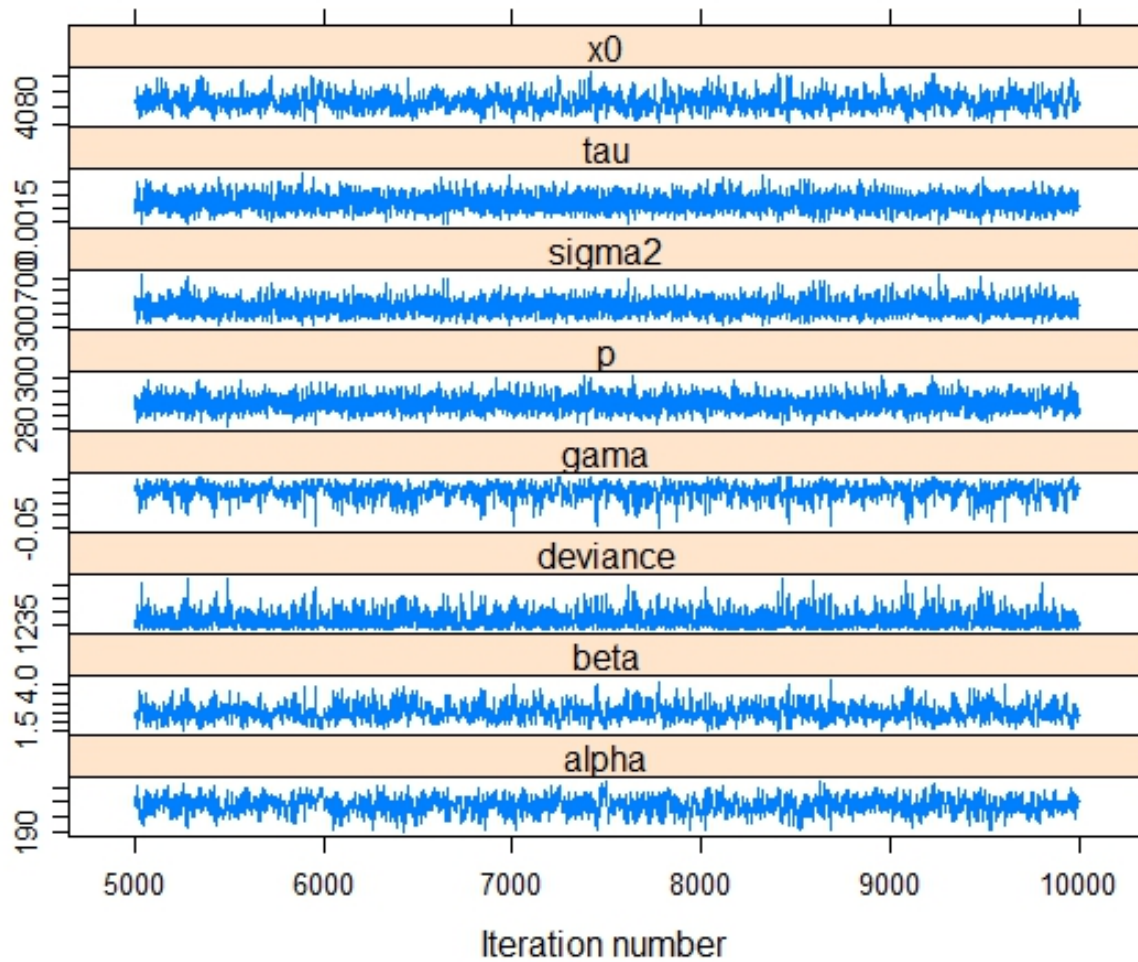


Figura 20 - Cadeias dos processos numéricos para obtenção das distribuições a posteriori para parâmetros do MQP com homogeneidade de variância.

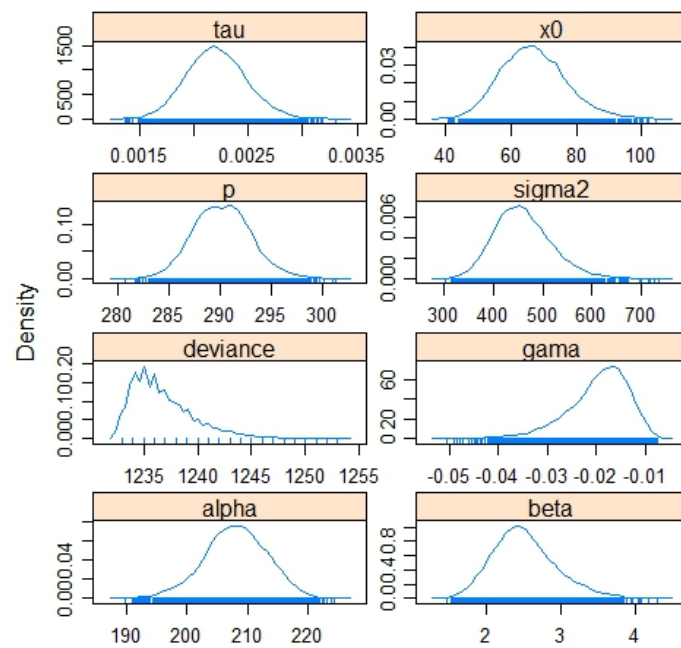


Figura 21 - Densidades a posteriori dos parâmetros do MQP com homogeneidade de variância.

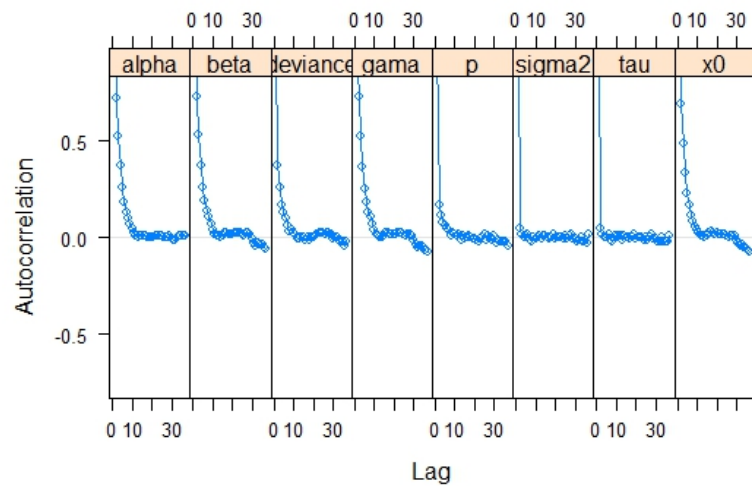


Figura 22 - ACF dos parâmetros do MQP com homogeneidade de variância.

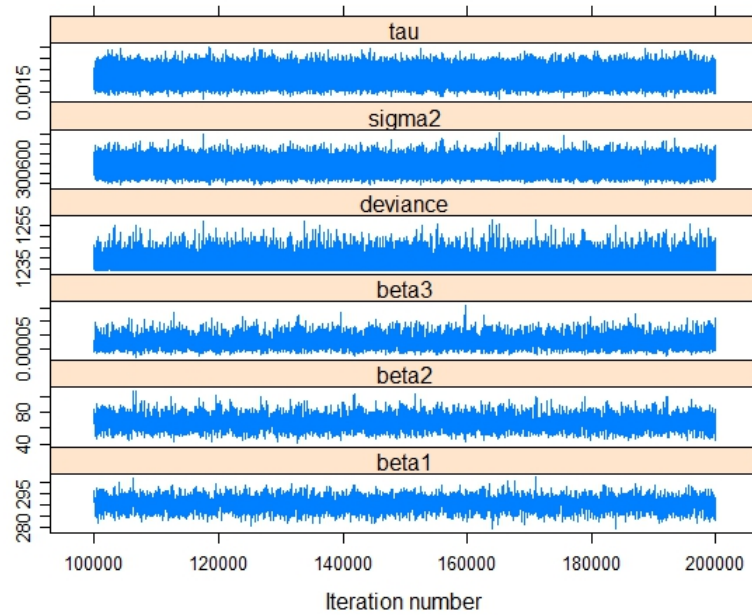


Figura 23 - Cadeias dos processos numéricos para obtenção das distribuições a posteriori para parâmetros do MNL com homogeneidade de variância.

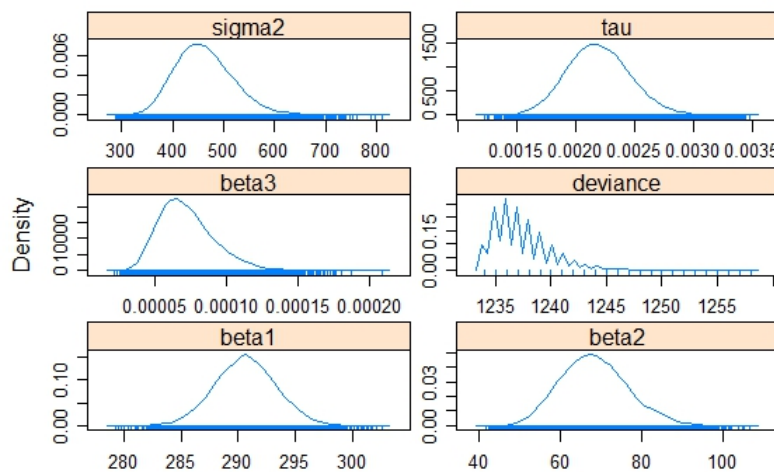


Figura 24 - Densidades a posteriori dos parâmetros do MNL com homogeneidade de variância.

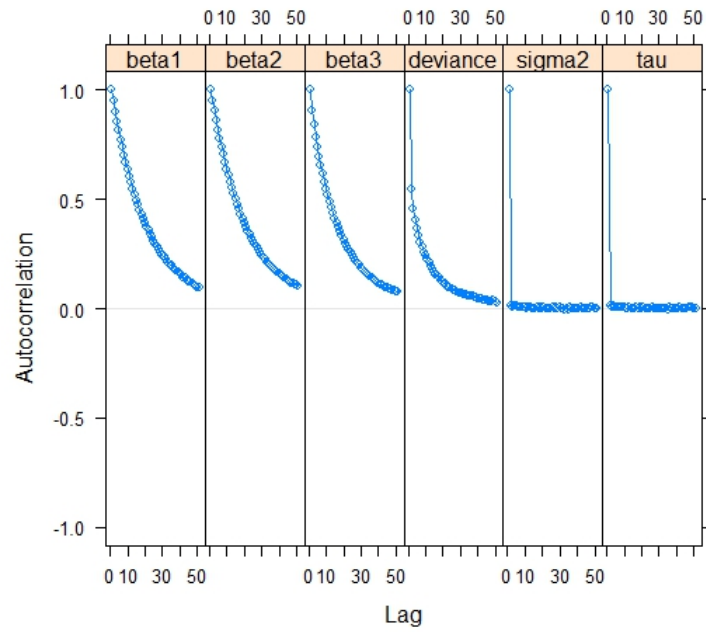


Figura 25 - ACF dos parâmetros do MNL com homogeneidade de variância.

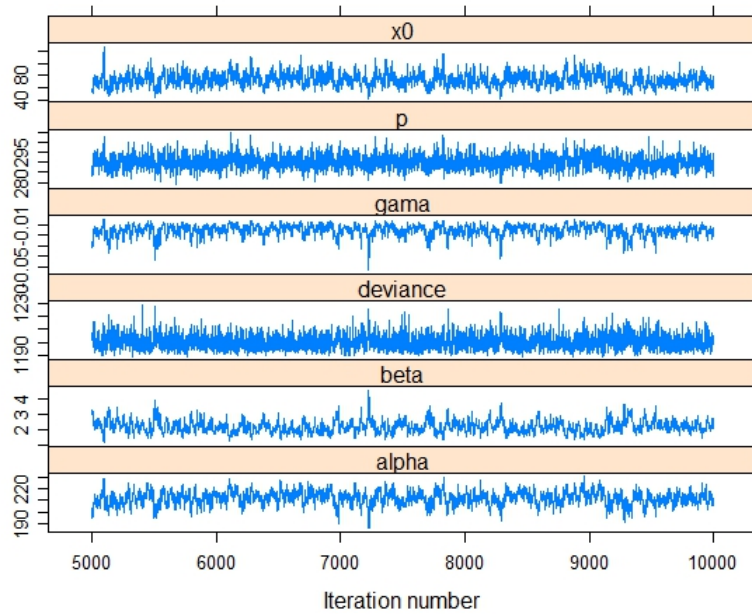


Figura 26 - Cadeias dos processos numéricos para obtenção das distribuições a posteriori para parâmetros do MQP com heterogeneidade de variância.



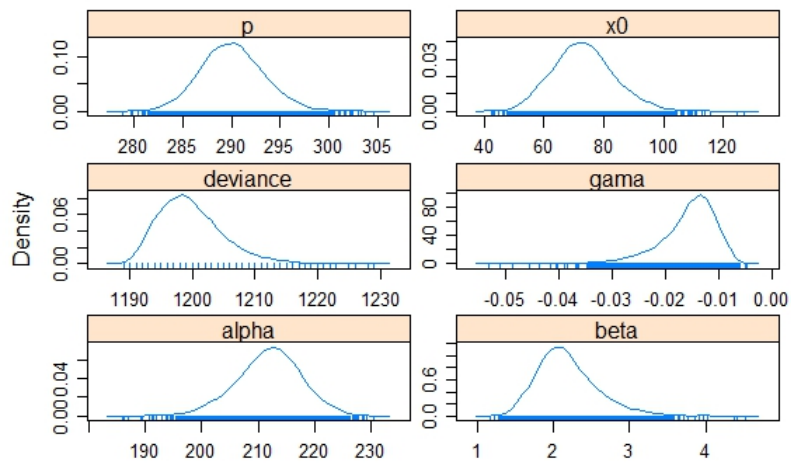


Figura 27 - Densidades a posteriori dos parâmetros do MQP com heterogeneidade de variância.

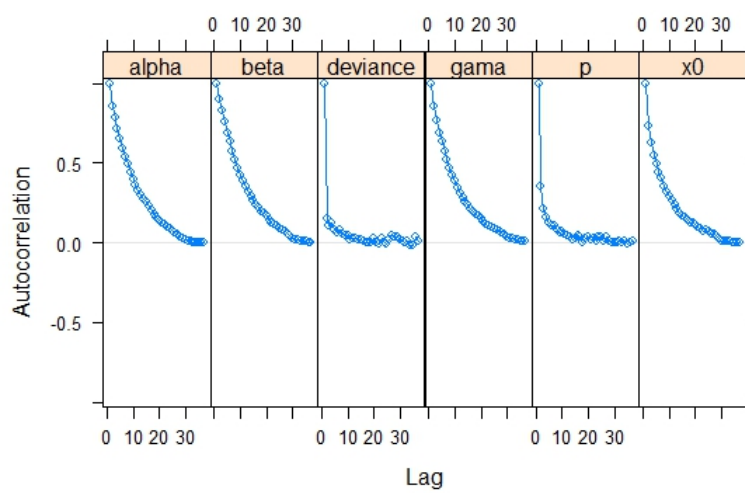


Figura 28 - ACF dos parâmetros do MQP com heterogeneidade de variância.

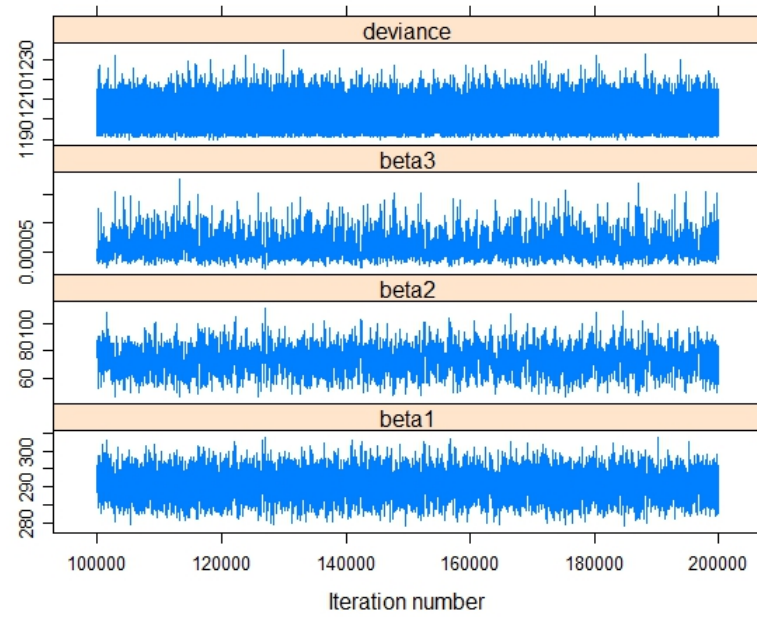


Figura 29 - Cadeias dos processos numéricos para obtenção das distribuições a posteriori para parâmetros do MNL com heterogeneidade de variância.

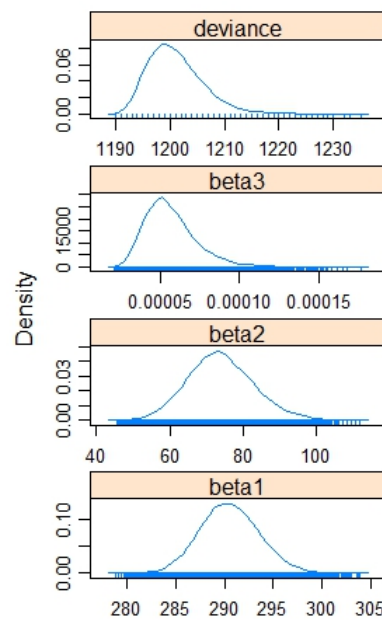


Figura 30 - Densidades a posteriori dos parâmetros do MNL com heterogeneidade de variância.

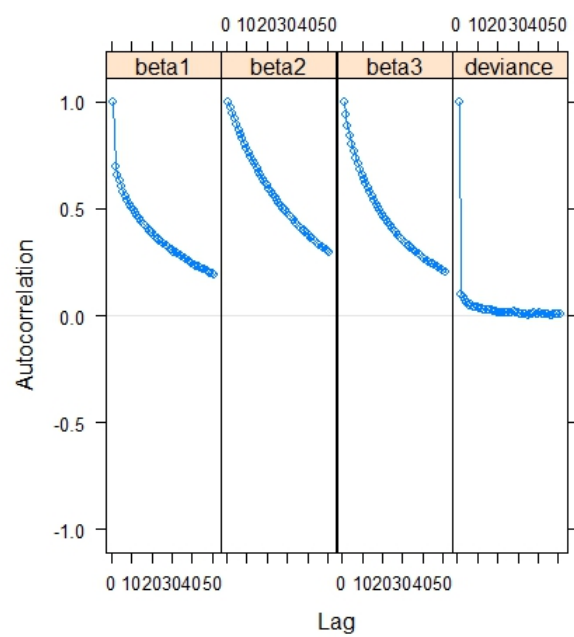


Figura 31 - ACF dos parâmetros do MNL com heterogeneidade de variância.