
**UNIVERSIDADE FEDERAL DE PERNAMBUCO**

**CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA**

**PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA**

**SUELEM TORRES DE FREITAS**

# ANÁLISE BAYESIANA DOS MODELOS DE REGRESSÃO LINEAR COM ERROS SIMÉTRICOS AUTORREGRESSIVOS E DADOS INCOMPLETOS

**Recife**

**2022**

**SUELEM TORRES DE FREITAS**

# ANÁLISE BAYESIANA DOS MODELOS DE REGRESSÃO LINEAR COM ERROS SIMÉTRICOS AUTORREGRESSIVOS E DADOS INCOMPLETOS

Dissertação apresentada ao Programa de Pós-Graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de mestre em Estatística. Área de Concentração: Estatística

Orientador: Prof. Dr. Aldo William Medina Garay

Co-Orientador: Prof. Dr. Rolando de la Cruz

**Recife**

**2022**

Catalogação na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

F866a Freitas, Suelem Torres de
Análise bayesiana dos modelos de regressão linear com erros simétricos autorregressivos e dados incompletos / Suelem Torres de Freitas. – 2022.
57 f.:il., fig, tab.

Orientador: Aldo William Medina Garay.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Estatística, Recife, 2022

Inclui referências e apêndice.

1. Estatística. 2. Modelo de regressão autoregressivo. I. Garay, Aldo William Medina (orientador). II. Título.

310 CDD (23. ed.) UFPE - CCEN 2022-184

SUELEM TORRES DE FREITAS

# ANÁLISE BAYESIANA DOS MODELOS DE REGRESSÃO LINEAR COM ERROS SIMÉTRICOS AUTORREGRESSIVOS E DADOS INCOMPLETOS

Dissertação apresentada ao Programa de Pós-Graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de mestre em Estatística. Área de Concentração: Estatística

Aprovada em: 29 de Agosto de 2022

BANCA EXAMINADORA

---

Prof. Dr. Aldo William Medina Garay (Presidente)
Universidade Federal de Pernambuco - UFPE

---

Prof. Dr. Celso Rômulo Barbosa Cabral
Universidade Federal do Amazonas - UFAM

---

Profa. Dra. Francyelle de Lima Medina
Universidade Federal de Pernambuco - UFPE

## RESUMO

Os modelos de regressão com erros autorregressivos considerando dados incompletos, isto é, quando a variável de interesse não está completamente disponível, seja pelo fato de ser censurada ou ausente, comumente denotada por *missing data*, tem se tornado um grande desafio para muitos pesquisadores. Existem técnicas de análise específicas para que a inferência com dados incompletos seja confiável, porém, apesar do crescente desenvolvimento de métodos nesta área, é recorrente encontrar o uso de inadequadas metodologias para a análise de dados incompletos. Uma suposição rotineira nestes tipos de modelos é considerar que as inovações seguem uma distribuição normal, no entanto, ao considerar a natureza sequencial dos dados analisados por este tipo de modelo, é conhecido que esta suposição pode não ser apropriada na presença de dados incompletos, assim este trabalho tem como objetivo principal apresentar uma abordagem Bayesiana dos modelos de regressão com erros autorregressivos, de ordem $p$, para dados incompletos (censurados ou *missing data*) supondo que as inovações seguem distribuições mais flexíveis, que possui como casos particulares as distribuições $t$ de Student, slash, normal contaminada e normal.



**Palavras-chave**: modelo de regressão autoregressivo; distribuições misturas de escala normal; dados incompletos; algoritmo MCMC.

## ABSTRACT

Regression models with autoregressive errors considering incomplete data, that is, when the variable of interest is not completely available, either due to being censored or absent, commonly denoted by missing data, has become a major challenge for many researchers. There are specific analysis techniques so that the inference with incomplete data is reliable, however, despite the growing development of methods in this area, it is recurrent to find the use of inadequate methodologies for the analysis of incomplete data. A routine assumption in these types of models is considering innovations following a normal distribution, however, when considering the nature sequence of the data analyzed by this type of model, it is known that this assumption may not be appropriate in the presence of incomplete data, so this work has as main objective to present a Bayesian approach to regression models with autoregressive errors, of order $p$, for incomplete data (censored or missingdata) assuming that the innovations follow more flexible distributions, which have the Student t, slash, contaminated normal and normal distributions as particular cases.



**Keywords**: Regression models with autoregressive errors; scale mixtures of normal distributions; incomplete data; algorithm MCMC.

## LISTA DE FIGURAS

## LISTA DE TABELAS

**SUMÁRIO**

# 1 INTRODUÇÃO

É comum em qualquer tipo de pesquisa a ocorrência de valores faltantes (*missing data*) e/ou valores censurados nas bases de dados. Nesse sentido, estes conjuntos de dados que estão incompletos, sejam com valores totalmente desconhecidos ou conhecidos parcialmente, podem ter diferentes efeitos na estatística inferencial. Segundo Wood, White e Thompson (2004), um conjunto de observações incompletas pode acarretar consequências, como por exemplo, viés substancial nas respostas, tendências, erros de má interpretação e redução na precisão de resultados.

Dados coletados ao longo do tempo vêm sendo estudados por meio da composição de dois modelos ou mais, tais como, os modelos de regressão linear com erros autoregressivos (OHTANI, 1990; CHIB, 1993; MCKNIGHT; MCKEAN; HUITEMA, 2000; ROSADI; FILZMOSER, 2019). Algumas características importantes deste tipo de dados é que as observações sequenciais em regra são dependentes, devendo ser levada em consideração sua ordem cronológica, frequência e variação, condições que geralmente devem ser examinadas concomitantemente. Além disso, nestas séries devido a natureza sequencial é complexo trabalhar com dados incompletos (EHLERS, 2007).

Existem pesquisas realizadas com conjunto de dados incompletos no intuito de enfrentar os problemas inerentes à presença de *missing data* e valores censurados na variável resposta em modelos de regressão linear com erros autoregressivos. Nesse contexto, vale ressaltar que metodologias para analisar esses impasses foram sugeridas, a exemplo dos estudos de Zeger e Brookmeyer (1986), Shin e Sarkar (1994), Thomson, Hossain e Ghahramani (2015), Schumacher, Lachos e Dey (2017) e Wang e Chan (2018).

Uma suposição rotineira em modelos de regressão com erros autoregressivos é considerar as inovações seguindo distribuição normal, no entanto, é conhecido que esta suposição pode não ser apropriada em especial na presença de *outliers* e/ou dados incompletos. Alguns autores estenderam para conjunto de dados completos modelos autoregressivos (AR) com inovações gaussianas para modelos AR com inovações seguindo uma distribuição de cauda pesada como a distribuição $t$ de Student (TIKU *et al.*, 2000; TARAMI; POURAHMADI, 2003; CHRISTMAS; EVERSON, 2010; HAGHBIN; NEMATOLLAHI, 2013; NDUKA, 2018).

A partir de então, estudos considerando modelos de regressão linear com erros AR com inovações seguindo uma distribuição $t$ de Student para dados incompletos na variável resposta foram realizados, como os estudos de Liu, Kumar e Palomar (2019), Zhou *et al.* (2021)

e Valeriano *et al.* (2021).

Liu, Kumar e Palomar (2019) observando os estudos de Tiku *et al.* (2000), Christmas e Everson (2010), afirmam que as metodologias para a estimativa de parâmetros para séries temporais AR de cauda pesada requerem dados completos. Dessa forma, portanto, não apresentam-se como as mais adequadas a serem utilizadas em cenários com dados faltantes e/ou censurados.

Uma classe de distribuições de cauda pesadas existentes para estimativas robustas é a classe de misturas da escala de distribuições normais (MEN) apresentadas por Andrews e Mallows (1974). Como essas distribuições têm caudas mais pesadas que a normal, em regra elas demonstram ser uma escolha mais interessante para análise inferencial com resultados robustos, visto que incluem, como casos especiais, muitas distribuições simétricas, como a normal, *t* de Student, slash e normal contaminada. Garay *et al.* (2015) e Garay *et al.* (2017) mostraram que os modelos MEN tanto em abordagens frequentistas, quanto Bayesianas, apresentam melhores resultados do que a distribuição normal na presença de valores censurados.

Pesquisas considerando modelos AR com inovações seguindo classes de distribuições de cauda pesada foram elaboradas como é o caso dos estudos de Maleki e Nematollahi (2017), Maleki *et al.* (2017), Ghasami, Khodadadi e Maleki (2020) e Ghasami, Maleki e Khodadadi (2020).

Existem técnicas de análise específicas para que a inferência com dados incompletos seja confiável, porém, apesar do crescente desenvolvimento de métodos nesta área, é recorrente encontrar o uso de inadequadas metodologias para a análise de dados incompletos. Como visto, embora existam trabalhos considerando modelos AR com inovações gaussianas ou inovações com distribuições de cauda pesada para dados incompletos, não há modelos de regressão linear com erros autoregressivos com inovações seguindo classes de distribuições de cauda pesada, uma vez que a estimativa de parâmetros nesse caso a princípio seria de difícil solução.

Ademais, constatamos que inexistem estudos que analisem, simultaneamente, tanto respostas censuradas quanto valores faltantes modelados pela classe de distribuição mistura de escala da normal, sendo esta considerada uma das famílias mais importante de distribuições. O objetivo do presente trabalho é abordar esse desafio e estudar o modelo de regressão linear autoregressivo de ordem $p$ com inovações seguindo uma distribuição mistura da escala normal para dados incompletos (RL-AR(p)-MEN-DI) sob uma abordagem Bayesiana.

Esta dissertação está organizada em seis capítulos. No Capítulo 2, apresentamos o marco teórico que contém as concepções que nortearam o trabalho, baseada nos estudos de

diversos autores e temas; no Capítulo 3 apresentamos o modelo proposto neste trabalho, o modelo de regressão autoregressivo com inovação mistura de escala da distribuição normal para dados incompletos. Em seguida, abordamos um estudo de inferência Bayesiano por meio da elaboração do amostrador de Gibbs para a estimação dos parâmetros do modelo; no Capítulo 4, apresentamos três estudos de simulação conduzidos com a finalidade de avaliar o desempenho das estimativas dos parâmetros do modelo, analisar a performance do modelo de regressão autoregressivo com inovação mistura de escala da distribuição normal para dados incompletos e verificar a vantagem de se trabalhar com uma distribuição de caudas pesadas, quando um conjunto de dados é incompleto; no Capítulo 5, apresentamos a aplicação do modelo proposto a dados reais e realizamos uma análise comparativa entre os modelos. Finalmente, no Capítulo 6, apresentamos as considerações finais e as perspectivas de trabalhos futuros.

# 2 MARCO TEÓRICO

## 2.1 NOTAÇÕES BÁSICAS

Nesta seção serão apresentadas algumas notações que serão utilizadas no presente trabalho. Denotamos um vetor aleatório por uma letra maiúscula e sua realização pela correspondente minúscula – como $\mathbf{Y}$ e $\mathbf{y}$, por exemplo. A transposta de uma matriz $\mathbf{A}$ denotamos por $\mathbf{A}^\top$. Seja $X \sim \mathrm{N}(\mu, \sigma^2)$ uma variável aleatória seguindo uma distribuição normal com média $\mu$ e variância $\sigma^2$, $\omega(\cdot|\mu, \sigma^2)$ denota a sua função de densidade de probabilidade (fdp) e $\Omega(\cdot)$ denota a função de distribuição acumulada (fda) da normal padrão, isto é, quando $\mu = 0$ e $\sigma^2 = 1$.

Para dois vetores aleatórios arbitrários $\mathbf{X}$ e $\mathbf{Y}$, $\mathbf{X} \overset{\mathrm{d}}{=} \mathbf{Y}$ denota que $\mathbf{X}$ tem a mesma distribuição que $\mathbf{Y}$ e $\mathbf{X} \perp \mathbf{Y}$ representa que eles são independentes. Se $\mathbf{Y}_i, i = 1, \ldots, n$ são vetores aleatórios, então $\mathbf{Y}_i \overset{\text{iid.}}{\sim} K$ caracteriza que todos os $\mathbf{Y}_i's$ são independentes e têm a mesma distribuição $K$.

Em seguida, definimos a distribuição truncada, seja $\mathbf{Y}$ um vetor aleatório e $\mathscr{B} \subset \mathbb{R}$ um conjunto de Borel, a distribuição condicional de $\mathbf{Y}|\{\mathbf{Y} \in \mathscr{B}\}$ é chamada de *distribuição de* $\boldsymbol{Y}$ *truncada em* $\mathscr{B}$. Assumindo que $\mathbf{Y}$ tem densidade $g(\cdot)$ e que $P\{\mathbf{Y} \in \mathscr{B}\} > 0$, a função densidade de $\mathbf{Y}|\{\mathbf{Y} \in \mathscr{B}\}$ é dada por

$$f(\mathbf{y}) = \frac{1}{P\{\mathbf{Y} \in \mathscr{B}\}} g(\mathbf{y}) \mathbb{I}_{\mathscr{B}}(\mathbf{y}), \tag{2.1}$$

em que $\mathbb{I}_{\mathscr{B}}(\cdot)$ denota a função indicadora, na qual, $\mathbb{I}_{\mathscr{B}}(\mathbf{y}) = 1$ se $\mathbf{y} \in \mathscr{B}$ e $\mathbb{I}_{\mathscr{B}}(\mathbf{y}) = 0$, caso contrário.

## 2.2 DADOS INCOMPLETOS

O problema com dados faltantes é relativamente comum em qualquer tipo de banco de dados, seja no âmbito acadêmico, empresarial ou industrial. Como informações não disponíveis, em um determinado banco de dados, tais dados ausentes poderiam afetar negativamente as conclusões extraíveis daquele conjunto de dados. Segundo Little e Rubin (2019, p.4) "dados ausentes são valores não observados que seriam significativos para análise se observados; em outras palavras, um valor ausente esconde um valor significativo".

Na visão de McKnight *et al.* (2007) pode-se dizer que os dados faltantes representam um obstáculo à segurança das análises científicas. Neste contexto, é presumível que, caso as informações acerca da natureza de um determinado evento sejam prejudicadas por quaisquer

dados faltantes, os pesquisadores buscariam a resolução desse problema. Para o mesmo autor, o mais preocupante, contudo, é constatar que há indícios práticos de que na verdade isso não ocorre. Assim, preocupa que a preponderância de dados faltantes e a insuficiente preocupação com eles afetem negativamente todo estudo científico realizado, isso porque indicam um efeito previsível na eficácia e interpretação dos resultados das pesquisas.

É interessante, aliás, destacar também a ocorrência dos chamados dados censurados, variáveis que não foram completamente observadas, os quais podem decorrer por várias razões, em alguns casos a censura ocorre quando a variável de interesse é o tempo até a ocorrência de um evento, nestes casos as censuras correspondem aos indivíduos que não experimentaram o evento de interesse antes do término do estudo, em outros, por exemplo, caso que utilizaremos neste trabalho, ocorre quando um instrumento de medição tem uma capacidade máxima fixa e que não fornece a quantia de interesse quando esta é ultrapassada, entre outros. Na opinião de Koul e Deshpande (1995), pode-se dizer que esses dados censurados manifestam-se espontaneamente em testes de vida, confiabilidade, trabalhos da área médica e experimentos clínicos.

Para Koul e Deshpande (1995), o exame de informações censuradas tem sido uma das prevalentes considerações dos estatísticos em seus estudos, como no caso dos estudos de Garay *et al.* (2015) e Garay *et al.* (2017). Mesmo assim, não parece haver razão para que a questão dos dados censurados seja tida como superada. Dessa forma, quaisquer tipos de dados incompletos, sejam ausentes ou censurados, em uma determinada análise estatística, devem ser tratados com muita parcimônia e atenção. Sendo, portanto, valiosos os esforços empregados para a mitigação dos efeitos adversos da falta desses dados.

## 2.3 MODELO AUTOREGRESSIVO DE ORDEM $p$

Seja o processo estocástico $\{x_t\}$, em que $t \in \mathbb{Z}$. O modelo autoregressivo é um modelo no qual a observação $x_t$ pode ser explicada como uma função das $p$ observações passadas. Assim, a estrutura autoregressiva é expressa por:

$$X_t = \sum_{j=1}^{p} \phi_j x_{t-j} + \eta_t, \tag{2.2}$$

sendo chamado modelo autoregressivo de ordem $p$, ou AR($p$), em que $X_t$ é estacionário, $\phi = (\phi_1, \ldots, \phi_p)^\top$ é o vetor de coeficientes autoregressivos e $\eta_t$ são variáveis aleatórias não correlacionadas e identicamente distribuídas com média 0 e variância $\sigma_\eta^2$, definidas como ruído

branco (SHUMWAY; STOFFER, 2000).

Seja $B$ o operador *backward* definido por $B^j x_t = x_{t-j}$. O modelo AR($p$) pode ser escrito por meio do operador autoregressivo de ordem $p$, $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$ da seguinte forma

$$\phi(B)x_t = \eta_t, \tag{2.3}$$

Considerando um processo estacionário, o processo pode ser escrito da forma

$$x_t = \psi(B)\eta_t,$$

em que $\psi(B) = \sum_{j=0}^{\infty} \psi B^j$ e $\phi(B)\psi(B) = 1$, ou seja,

$$(1 - \phi B - \phi_2 B^2 - \cdots - \phi_p B^p)(1 + \psi_1 B + \psi_2 B^2 + \cdots + \psi_j B^j + \ldots) = 1. \tag{2.4}$$

Assim, para diferentes valores de $p$, é possível obter os valores dos coeficientes $\psi_j's$ por meio da solução da Equação (2.4), tal que:

- Para $p = 1$

$$\psi_j = \phi_1^j \tag{2.5}$$

- Para $p \geq 2$, quando $j \geq p$

$$\psi_j = \psi_{j-1}\phi_1 + \psi_{j-2}\phi_2 + \cdots + \psi_{j-p}\phi_p, \tag{2.6}$$

  com $\psi_0 = 1$.

## 2.4 DISTRIBUIÇÕES MISTURAS DA ESCALA NORMAL

Andrews e Mallows (1974) apresentaram uma família de distribuições simétricas chamada "distribuições de mistura de escala normal". A vantagem desta classe é incorporar um parâmetro de forma (relacionado à curtose e não à assimetria) à densidade normal e, dessa forma, obter distribuições mais flexiveis e com caudas mais pesadas do que a distribuição normal, bastante úteis em inferência robusta para dados simétricos.

Assim, uma variável aleatória $Y$ tem uma distribuição mistura da escala normal, com parâmetro de locação $\mu$ e parâmetro de escala $\sigma^2 > 0$, a qual denotamos por $Y \sim \text{MEN}(\mu, \sigma^2, \nu)$, se tem a seguinte representação estocástica:

$$Y = \mu + U^{-\frac{1}{2}} Z; \quad Z \sim N(0, \sigma^2); \quad Z \perp U, \tag{2.7}$$

em que $U$ é uma variável aleatória positiva, conhecida como fator escala, com fda $H(\cdot|\nu)$, o qual é chamada de função de distribuição de mistura e $\nu$ é um escalar ou vetor de parâmetros indexado à distribuição de U.

Quando $\mu = 0$ e $\sigma^2 = 1$ temos a distribuição MEN padronizada. Nesse caso, denotamos a função de distribuição por $F_{\text{MEN}}(\cdot)$. A variável aleatória $U$ pode ser discreta ou contínua e a forma da distribuição MEN é determinada pela distribuição do fator escala $U$. Desse modo, apresentamos os casos utilizados neste trabalho:

- *Distribuição Normal*: neste caso o fator de escala $U$ é uma variável aleatória degenerada em 1, isto é, $P(U = 1) = 1$.
- *Distribuição t de Student*: neste caso, $U \sim \text{Gamma}(\nu/2, \nu/2)$, com $\nu > 0$, em que $\text{Gamma}(a,b)$ denota a distribuição Gama com média $a/b$. A fdp da variável aleatória $Y$, definida na Equação (2.7) é dada por

  $$f_{\text{TS}}(y|\nu) = \frac{1}{B(\nu/2, 1/2)\sqrt{\nu}} \left(1 + \frac{y^2}{\nu}\right)^{-(\nu+1)/2},$$

  em que $\nu > 0$ é parâmetro de forma e $B(a,b)$ representa a função beta. Utilizamos a notação $Y \sim \text{TS}(\mu, \sigma^2; \nu)$.

  A distribuição t de Student converge ao caso normal quando $\nu \to \infty$.
- *Distribuição Slash*: neste caso, a distribuição do fator de escala $U$ é $\text{Beta}(\nu, 1)$, com $\nu > 0$. A função de densidade da variável aleatória $Y$, definida na Equação (2.7), é dada por

  $$f_{\text{SL}}(y|\nu) = \nu \int_0^1 u^{\nu-1} \omega(y\sqrt{u}) du, \quad y \in \mathbb{R}.$$

  Utilizamos a notação $Y \sim \text{SL}(\mu, \sigma^2; \nu)$.
- *Distribuição Normal Contaminada*: neste caso, $U$ é uma variável aleatória discreta que pode assumir dois valores, 1 ou $\delta$. Dessa forma, a função de probabilidade de $U$ é dada por

  $$h(u|\nu, \delta) = \nu \mathbb{I}_{\{\delta\}}(u) + (1 - \nu)\mathbb{I}_{\{1\}}(u)\nu, \quad \delta \in (0,1).$$

  Segue imediatamente que a densidade da variável aleatória $Y$, definida na Equação (2.7), é dada por

  $$f_{\text{NC}}(y|\nu, \delta) = \nu\omega(y|0, \delta^{-1}) + (1 - \nu)\omega(y).$$

  Utilizamos a notação $Y \sim \text{NC}(\mu, \sigma^2; \nu)$.

Considere uma variável aleatória $Y$ com uma distribuição MEN truncada no intervalo $\lfloor a,b \rfloor$, para b > a, ou seja, $Y \stackrel{\mathrm{d}}{=} X|(X \in \lfloor a,b \rfloor)$, em que $X \sim \mathrm{MEN}(\mu,\sigma^2,\nu)$. Neste caso, $-\infty \leqslant a < b \leqslant \infty$ e a notação $\lfloor a,b \rfloor$ denota que o intervalo pode ser aberto, fechado ou semi-aberto. Neste caso, de acordo com a Equação (2.1), a variável aleatória $Y$ tem função de densidade

$$f_{\mathrm{MENT}}(y|\mu,\sigma^2,\nu;\lfloor a,b \rfloor) = \frac{f_{\mathrm{MEN}}(y|\mu,\sigma^2,\nu)}{\left[F_{\mathrm{MEN}}\left(\frac{b-\mu}{\sigma}\right) - F_{\mathrm{MEN}}\left(\frac{a-\mu}{\sigma}\right)\right]}\mathbb{I}_{\lfloor a,b \rfloor}(y),$$

em que $F_{\mathrm{MEN}}$ é a fda da distribuição MEN padronizada. Utilizamos a notação $Y \sim \mathrm{MENT}(\mu,\sigma^2, \nu;\lfloor a,b \rfloor)$.

## 2.5 INFERÊNCIA BAYESIANA

Na abordagem Bayesiana assume-se que os parâmetros possuem uma distribuição de probabilidade, este tipo de abordagem é desenvolvida na presença de um conjunto de observações $y = (y_1,\ldots,y_n)$, cujo valores são descritos por uma densidade ou função de probabilidade $f(y|\theta)$, onde $\theta$ são os parâmetros da distribuição. Para mais detalhes, ver os livros utilizados neste trabalho (BOX; TIAO, 1992; ALBERT, 2007; RIZZO, 2007; NTZOUFRAS, 2009; CASELLA; BERGER, 2010).

### 2.5.1 Distribuição a priori e a posteriori

Quando se estuda uma população, é possível que o pesquisador tenha algum conhecimento sobre o parâmetro de interesse $\theta$. A abordagem Bayesiana considera que este conhecimento possa ser formalmente incorporado na análise. Esta informação prévia pode ser inserida na análise por meio de uma densidade ou função de probabilidade $\pi(\theta)$, chamada de distribuição a priori.

Na inferência Bayesiana, temos dois componentes para estimar os parâmetros: a distribuição da amostra $f(y|\theta)$ e a distribuição $\pi(\theta)$. A distribuição amostral fornece a função de verossimilhança $L(\theta|y) = f(y_1|\theta)\cdots f(y_n|\theta)$. Com base nisso, pode-se encontrar a distribuição de $\theta$ após observar $y$. Esta distribuição é chamada de distribuição a posteriori $\pi(\theta|y)$, tal que

$$\pi(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{\int f(y|\theta)\pi(\theta)d\theta},$$

em que o denominador é uma constante em relação a $\theta$. A distribuição a posteriori pode ser reescrita como

$$\pi(\theta|y) = \frac{L(\theta|y)\pi(\theta)}{h(y)},$$

em que $h(y)$ é denominada constante normalizadora e representa a densidade marginal de $y$, podendo ser obtida da seguinte forma

$$h(y) = \int \pi(y,\theta)d\theta = \int L(\theta|y)\pi(\theta)d\theta.$$

### 2.5.2 Monte Carlo via Cadeias de Markov (MCMC)

Em inferência Bayesiana, haja vista ser rotineiro que a forma de distribuição a posteriori não resulte em uma distribuição de probabilidade conhecida, sói utilizar-se uma técnica chamada Monte Carlo via Cadeias de Markov (MCMC). Em regra, esse evento se dá quando há apenas um único parâmetro, sendo mais observado quando $\theta$ é um vetor de $k$ parâmetros. O ponto chave é obter estimativas dos parâmetros desta distribuição com base na distribuição a posteriori, para detalhes, veja Gilks, Richardson e Spiegelhalter (1996).

As técnicas de MCMC baseiam-se na construção de uma cadeia de Markov, do qual resulta uma distribuição alvo (chamada estacionária ou de equilíbrio) que, no nosso caso, é a distribuição a posteriori $\pi(\theta|y)$.

Uma cadeia de Markov de tempo discreto que evoluem em espaços de estados finitos, por exemplo, é um processo estocástico $\{\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(T)}\}$ tal que

$$f(\theta^{(t+1)}|\theta^{(t)}, \ldots, \theta^{(1)}) = f(\theta^{(t+1)}|\theta^{(t)})$$

isto é, $\theta^{(t+1)}$ depende apenas do valor de $\theta^{(t)}$. Assim, quando a cadeia de Markov é recorrente positiva, ou seja, partindo de um estado $t$, o tempo esperado para o processo visitar novamente este estado é finito, irredutível, no qual todos os estados do processo se comunicam e aperiódica, apresentando intervalos irregulares para alcançar o estado, com $t \to \infty$, a distribuição de $\theta^{(t)}$ converge para sua distribuição de equilíbrio, o qual é independente dos valores iniciais da cadeia $\theta^{(0)}$; para detalhes, veja Gilks, Richardson e Spiegelhalter (1996).

Os métodos MCMC incorporam a noção de um procedimento iterativo (razão pela qual eles são frequentemente chamados de métodos iterativos), uma vez que em cada etapa eles produzem valores dependendo do anterior. Os algoritmos mais aplicados são o algoritmo de Metropolis-Hastings e o amostrador de Gibbs.

#### 2.5.2.1 O algoritmo de Metropolis-Hastings

A origem desse algoritmo se deu com Metropolis *et al.* (1953) que apresentando os métodos de simulação baseados em cadeias de Markov empregados na ciência, formularam o algoritmo Metropolis. Anos depois Hastings (1970) generalizou o método original no que é conhecido como algoritmo de Metropolis-Hastings. Nestes algoritmos, um valor é gerado de uma distribuição conhecida, sendo este valor aceito também como um ponto da distribuição a posteriori $\pi(\theta|y)$ com uma probabilidade $\alpha$. Este mecanismo de correção garante a convergência da cadeia para a distribuição de equilíbrio.

Suponha que a cadeia esteja no estado $\theta$ e um valor $\theta'$ é gerado de uma distribuição conhecida $q(\cdot|\theta)$. Esta distribuição é conhecida como distribuição proposta e ela geralmente depende do estado atual da cadeia. O novo valor $\theta'$ é aceito com probabilidade

$$\alpha(\theta,\theta') = \min\left\{\frac{\pi(\theta'|y)q(\theta|\theta')}{\pi(\theta|y)q(\theta'|\theta)},1\right\}.$$

Em termos práticos, o algoritmo de Metropolis-Hastings pode ser descrito pelos seguintes passos:

1. Especifique o valor inicial $\theta^0$;
2. Gere um novo valor $\theta'$ da distribuição $q(\cdot|\theta)$;
3. Calcule a probabilidade de aceitação $\alpha(\theta,\theta')$ e gere $u \sim \mathrm{U}(0,1)$;
4. Se $u \leq \alpha$ então aceite o novo valor e faça $\theta^{(t+1)} = \theta'$. Caso contrário rejeite e faça $\theta^{(t+1)} = \theta^{(t)}$;
5. Volte ao passo 2.

Uma característica importante do algoritmo é que não precisamos avaliar a constante normalizadora $h(y)$ envolvida em $\pi(\theta|y)$, uma vez que em $\alpha$ a constante é cancelada. Após a sequência de amostras retornadas, uma prática comum e bastante eficiente é realizar o *burn-in*, que consiste em eliminar os valores resultantes das primeiras iterações do algoritmo, realizadas antes de a convergência ser atingida. Além disso, existe uma dependência entre observações sucessivas, para eliminar ou minimizar esta correlação, o que se faz é guardar as observações espaçadas utilizando um passo constante, processo conhecido como *thinning*. Para mais detalhes ver Gilks, Richardson e Spiegelhalter (1996), Gamerman e Lopes (2006).

#### 2.5.2.2 Amostrador de Gibbs

O amostrador de Gibbs foi introduzido por Geman e Geman (1984) e é um caso especial da versão "componente-a-componente" do algoritmo de Metropolis-Hastings. Embora o amostrador de Gibbs seja um caso especial do algoritmo Metropolis-Hasting, é geralmente citada como uma técnica de simulação separada devido à sua popularidade e conveniência.

Considere $\theta = (\theta_1, \theta_2, \ldots, \theta_d)$ um vetor com $d$ dimensões. Este algoritmo consiste em amostrar um parâmetro por vez, baseado na sua distribuição condicional completa $\pi(\theta|\theta_{-i})$, onde $\theta_{-i} = (\theta_1, \theta_2, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_d)$. Para obter a distribuição condicional completa de $\theta_i$, determina-se os termos da distribuição conjunta a posteriori, apenas os termos que dependem de $\theta_i$ e verifica-se a proporcional a esta distribuição, observando se possui uma densidade conhecida que pode ser gerada facilmente.

O algoritmo do amostrador de Gibbs é dado da seguinte maneira:

1. Inicializar $\theta = (\theta_1^{(0)}, \ldots, \theta_d^{(0)})$;
2. Simular $\theta_1^{(1)}$ da condicional $\theta_1|\theta_2^{(0)}, \ldots, \theta_d^{(0)}$;
3. Simular $\theta_2^{(1)}$ da condicional $\theta_2|\theta_1^{(1)}, \theta_3^{(0)}, \ldots, \theta_d^{(0)}$;
4. Simular $\theta_d^{(1)}$ da condicional $\theta_d|\theta_1^{(1)}, \ldots, \theta_{d-1}^{(0)}$;
5. Voltar ao passo 2.

Assim, após um período de *burn-in*, $(\theta_1^{(k)}, \ldots, \theta_d^{(k)}), \ldots, (\theta_1^{(n)}, \ldots, \theta_d^{(n)})$ são realizações da distribuição de interesse.

Além destes algoritmos, outras variações podem ser propostas no MCMC. Por exemplo, fazer o algoritmo de Gibbs por passos de Metropolis, em que a distribuição condicional completa de cada parâmetro é a distribuição $\pi(y)$ do algoritmo de Metropolis-Hastings, e Metropolis por blocos, onde blocos $(\theta_1, \ldots, \theta_{k1}), \ldots, (\theta_{kp-1}, \ldots, \theta_{kp})$ são amostrados um por vez pelo algoritmo de Metropolis-Hastings.

No capítulo seguinte, descreveremos o modelo de regressão autoregressivo com inovações que seguem uma distribuição mistura de escala da distribuição normal a partir de uma abordagem Bayesiana, utilizando as definições e ferramentas descritas neste capítulo.

# 3 MODELO DE REGRESSÃO AUTOREGRESSIVO COM INOVAÇÕES MEN

## 3.1 DESCRIÇÃO DO MODELO

Considere um modelo de regressão linear autoregressivo de ordem $p$ com inovação seguindo uma distribuição mistura de escala normal (RL-AR($p$)-MEN). Assim, a representação deste modelo de regressão para as respostas observadas no tempo $t$ pode ser descrita como

$$\begin{aligned} Y_t &= \mathbf{x}_t^\top \beta + \varepsilon_t, \qquad t = 1,2,\ldots,T, \\ \varepsilon_t &= \sum_{i=1}^{p} \phi_i \varepsilon_{t-i} + \eta_t, \quad \eta_t \sim \text{MEN}(0,\sigma^2,\nu). \end{aligned} \tag{3.1}$$

Uma maneira alternativa de escrever o modelo (3.1) é dada por

$$Y_t = \mathbf{x}_t^\top \beta + \sum_{i=1}^{p} \phi_i (Y_{t-i} - \mathbf{x}_{t-i}^\top \beta) + \eta_t. \tag{3.2}$$

Para facilitar adotamos que $\mu_t^* = \mathbf{x}_t^\top \beta + \sum_{i=1}^{p} \phi_i (Y_{t-i} - \mathbf{x}_{t-i}^\top \beta)$.

Seguindo Garay *et al.* (2015) assumimos que a resposta $Y_t$ não é totalmente observada para todas as $t$ observações e consideramos censura à esquerda, no entanto, os resultados são facilmente estendíveis para outros tipos de censura. Sejam os $t$ dados observados, em que $V_t$ representa o valor da censura ou o valor sem censura e $C_t$ são os indicadores de censura, essas observações são expressas de tal forma que

$$Y_t \leqslant V_t \quad \text{se} \quad C_t = 1 \quad \text{e} \quad Y_t = V_t \quad \text{se} \quad C_t = 0. \tag{3.3}$$

Por outro lado, seguindo Liu, Kumar e Palomar (2019) e Zhou *et al.* (2021), também consideramos que alguns $Y_t$'s são desconhecidos devido a várias razões, os quais são indicados por *NA* (*missing values*). Estes valores faltantes aparecem dispostos no conjunto de dados de forma consecutiva ou sendo um único valor faltante entre duas observações, e, portanto, definidos como bloco de *NAs*.

No sentido de adquirir critérios suficientes para que a verossimilhança exata e inferências Bayesianas sejam válidas, eximindo a necessidade de modelar o mecanismo gerador de ausência de dados, assim como em Liu, Kumar e Palomar (2019) e Zhou *et al.* (2021), também assumimos que o mecanismo de dados faltantes é ignorável, para mais detalhes, ver Little e Rubin (2019).

Vamos supor que temos um conjunto de dados com D blocos faltantes, conforme Figura 1, em que no $d$-ésimo bloco faltante, existem $n_d$ observações faltantes $y_{t_d+1},\ldots,y_{t_d+n_d}$,

$$y_1,\ldots,y_{t_1},\overbrace{\text{NA},\ldots,\text{NA}}^{n_1},y_{t_1+n_1+1},\ldots,y_{t_d},\overbrace{\text{NA},\ldots,\text{NA}}^{n_d},$$
$$y_{t_d+n_d+1},\ldots,y_{t_D},\underbrace{\text{NA},\ldots,\text{NA}}_{n_D},y_{t_D+n_D+1},\ldots,y_T.$$

**Figura 1 – Conjunto de dados com valores faltantes.**

que são limitados na esquerda e na direita por dois dados observados $y_{t_d}$ e $y_{t_d+n_d+1}$, respectivamente. Denotamos o conjunto dos índices dos valores observados por $C_o$, o conjunto dos índices dos valores censurados por $C_c$ e o conjunto dos índices dos valores faltantes por $C_m$. Também denotamos $\mathbf{y} = (y_t, p+1 \leqslant t \leqslant T)$, $\mathbf{y}_o = (y_t, t \in C_o)$, $\mathbf{y}_c = (y_t, t \in C_c)$ e $\mathbf{y}_m = (y_t, t \in C_m)$.

Portanto, seja $\theta = (\beta^\top, \phi, \sigma^2, \nu)^\top$ o vetor dos parâmetros do modelo de regressão linear autoregressivo com inovação misturas de escala normal para dados incompletos, denotado como RL-AR($p$)-MEN-DI e seja $\mathbf{x} = (\mathbf{x}_1^\top, \ldots, \mathbf{x}_{T+p}^\top)$ as covariáveis, a função de verossimilhança de $\theta$ será definida mais adiante.

O modelo RL-AR($p$)-MEN-DI por se tratar de uma classe de distribuições simétricas, devido a família de distribuições de mistura de escala normal, daremos atenção especial a alguns de seus casos particulares como a distribuição normal, a t de Student, a slash e a normal contaminada, descritas na Seção (2.4) e denotadas como RL-AR($p$)-N-DI, RL-AR($p$)-T-DI, RL-AR($p$)-SL-DI e RL-AR($p$)-NC-DI, respectivamente, ao assumirem a inovação do modelo.

## 3.2 FUNÇÃO DE VEROSSIMILHANÇA

Como pode ser visto em Hamilton (1994) e Prado e West (2010), a função de verossimilhança exata requer o conhecimento da distribuição conjunta das primeiras $p$ observações e o procedimento de maximização pode resultar em um sistema de equações em $\theta$ e $(y_1,\ldots,y_{T+p})$, para o qual não há solução para $\theta$. Assim, considerando a presença de valores faltantes $\mathbf{y}_m$, uma alternativa para a maximização da função de verossimilhança exata é considerar a função de verossimilhança do dados observados $\mathbf{y}_{obs} = (\mathbf{y}_o, \mathbf{y}_c)$ condicionada as primeiras $p$ observações, dada por:

$$\begin{aligned} L(\theta|\mathbf{y}_{obs}) &\propto \int f(\mathbf{y}_o,\mathbf{y}_c,\mathbf{y}_m|\theta)\mathrm{d}\mathbf{y}_m \\ &\propto \left\{\int \prod_{t=p+1}^{p+T} f\left(y_t|y_{t-1},\ldots,y_{t-p},\theta\right)\right\}\mathrm{d}\mathbf{y}_m \\ &\propto \left\{\int \prod_{t=p+1}^{p+T} f_{\text{MEN}}\left(y_t|\mu_t^*,\sigma^2;\nu\right)^{1-\mathbb{I}_{\mathbf{y}_c}(y_t)} F_{\text{MEN}}\left(\frac{V_t-\mu_t^*}{\sigma}\right)^{\mathbb{I}_{\mathbf{y}_c}(y_t)}\right\}\mathrm{d}\mathbf{y}_m, \end{aligned} \tag{3.4}$$

em que $\{Y_t|y_{t-1},\ldots,y_{t-p},\theta\}$ segue uma distribuição MEN, com parâmetro de locação $\mu_t^* = \mathbf{x}_t^\top\beta + \sum_{i=1}^{p}\phi_i\varepsilon_{t-i}$ e parâmetro de escala $\sigma^2$, em que $\varepsilon_{t-i} = y_{t-i} - \mathbf{x}_{t-i}^\top\beta$ para $i = 1,\ldots,p$. $\mathbb{I}_{\mathscr{B}}(\cdot)$ denota a função indicadora, na qual, $\mathbb{I}_{\mathscr{B}}(y) = 1$ se $y \in \mathscr{B}$ e $\mathbb{I}_{\mathscr{B}}(y) = 0$, caso contrário e $F_{\text{MEN}}(\cdot)$ denota a fda da distribuição MEN padronizada, $\text{MEN}(0,1;\nu)$.

A seguir, apresentamos dois resultados importantes que foram desenvolvidos por Liu, Kumar e Palomar (2019, Lemas 1-2) relacionados ao conjunto de $M$ blocos de valores faltantes (B) $\mathbf{y}_m = (\mathbf{y}_{\text{B}1},\ldots,\mathbf{y}_{\text{B}M})$, com $\mathbf{y}_{\text{B}k} = \{y_{s_k+1},\ldots,y_{s_k+m_k}\}$, no contexto do modelo RL-AR($p$)-N-DI, que será crucial para desenvolver nosso algoritmo MCMC.

(1) Dado $(\mathbf{y}_{obs},\theta)$, os blocos de valores faltantes $\mathbf{y}_m$, são independentes, ou seja,

$$f(\mathbf{y}_m|\mathbf{y}_{obs},\theta) = \prod_{k=1}^{M} f(\mathbf{y}_{\text{B}k}|\mathbf{y}_{obs},\theta).$$

(2) A distribuição do $k$-ésimo bloco de valores faltantes $\mathbf{y}_{\text{B}k}$, dado $(\mathbf{y}_{obs},\theta)$, depende apenas das $p$ observações anteriores ao bloco $\mathbf{y}_{(s_k-p)} = \{y_{s_k-1},\ldots,y_{s_k-p}\}$ e das $p$ observações posteriores ao bloco $\mathbf{y}_{(s_k+m_k+p)} = \{y_{s_k+m_k+1},\ldots,y_{s_k+m_k+p}\}$.

Embora este último resultado tenha sido obtido anteriormente por Liu, Kumar e Palomar (2019), fornecemos uma prova alternativa e, em seguida, apresentamos uma forma fechada para o modelo autoregressivo de ordem $p \in \{1,2,3\}$.

Para obter a distribuição de $\mathbf{y}_{\text{B}k}$ dado $\big(\mathbf{y}_{(s_k-p)},\mathbf{y}_{(s_k+m_k+p)}\big)$, denotado por $f\big(\mathbf{y}_{\text{B}k}|\mathbf{y}_{(s_k-p)}, \mathbf{y}_{(s_k+m_k+p)},\theta\big)$, é necessário analisar a distribuição conjunta de $\mathbf{y}_{\text{B}k}$ e $\mathbf{y}_{(s_k+m_k+p)}$, dado $\big(\mathbf{y}_{(s_k-p)},\theta\big)$. Sob condições de estacionariedade, cada observação do vetor $\mathbf{y}_{kp} = \big(\mathbf{y}_{\text{B}k},\mathbf{y}_{(s_k+m_k+p)}\big)$ pode ser escrita usando iterações de retrogressão $j$ vezes, para $j = 1,2,\ldots,m_k,m_k+1,\ldots,m_k+p$, com

$$Y_{s_k+j} = \mu_{s_k+j} + \sum_{r=0}^{j-1}\psi_r\eta_{s_k+j-r},$$

em que $\mu_{s_k+j}$ são funções de $\varepsilon_{(s_k-p)} = \{\varepsilon_{s_k-1},\ldots,\varepsilon_{s_k-p}\}$, $\mathbf{x}_{s_k+j}^\top\beta$ e $\varepsilon_{s_k} = Y_{s_k} - x_{s_k}^\top\beta$. Os coeficientes $\psi_r$ são funções de $\phi = (\phi_1,\ldots,\phi_p)^\top$. Observe que $\mu_{s_k+j}$ e $\psi_r$ dependem da $p$-ésima ordem do modelo autoregressivo, ou seja:

- Para $p = 1$:

  $\mu_{s_k+j} = \mathbf{x}_{s_k+j}^\top\beta + \psi_j\varepsilon_{s_k}$,

  $\psi_r = \phi^r$.

- Para $p = 2$:

  $\mu_{s_k+j} = \mathbf{x}_{s_k+j}^\top\beta + \psi_j\varepsilon_{s_k} + \phi_2\psi_{j-1}\varepsilon_{s_k-1}$,

  $\psi_0 = 1$; $\psi_1 = \phi_1$,

  $\psi_r = \phi_1\psi_{r-1} + \phi_2\psi_{r-2}$, para todo $r \geq 2$.

- Para $p = 3$:

  $\mu_{s_k+j} = \mathbf{x}_{s_k+j}^{\top}\beta + \psi_j \varepsilon_{s_k} + (\phi_2\psi_{j-1} + \phi_3\psi_{j-2})\varepsilon_{s_k-1} + \phi_3\psi_{j-1}\varepsilon_{s_k-2}$,

  $\psi_0 = 1$; $\psi_1 = \phi_1$; $\psi_2 = \phi_1^2 + \phi_2$,

  $\psi_r = \phi_1\psi_{r-1} + \phi_2\psi_{r-2} + \phi_3\psi_{r-3}$, para todo $r \geq 3$.

Dessa forma, $Y_{s_k+j}$, para todo $j$, pode ser representado como a soma das constantes $\mu_{s_k+j}$ e do processo linear $\sum_{r=0}^{j-1}\psi_r\eta_{s_k+j-r}$, com $\eta_t \overset{\text{iid.}}{\sim} \mathrm{N}(0,\sigma^2)$. Como consequência, dado as observações anteriores $\mathbf{y}_{(s_k-p)}$, nós temos que:

$$\mathbf{y}_{kp}|\mathbf{y}_{(s_k-p)}, \theta \sim \mathrm{N}_{m_k+p}(\mu_{kp}, \Sigma_{kp}), \tag{3.5}$$

em que $\mu_{kp_{(j)}} = \mu_{s_k+j}$ e

$$\Sigma_{kp_{(i,j)}} = Cov\left\{\mu_{s_k+i} + \sum_{r=0}^{i-1}\psi_r\eta_{s_k+i-r}, \mu_{s_k+j} + \sum_{r=0}^{j-1}\psi_r\eta_{s_k+j-r}\right\} = \sum_{r=0}^{\min\{i-1,j-1\}}\psi_i\psi_{|i-j|+r}\sigma^2.$$

## 3.3 INFERÊNCIA BAYESIANA PARA O MODELO RL-AR($p$)-MEN-DI

### 3.3.1 Distribuições a priori

Seja $\theta = (\beta, \phi, \sigma^2, \nu)^{\top}$ o vetor dos parâmetros do modelo RL-AR($p$)-MEN-DI, propomos as seguintes distribuições prioris:

(i) $\beta \sim \mathrm{N}_q(\mu_\beta, \Sigma_\beta)$, em que $\mu_\beta$ é o vetor $q \times 1$ de hiperparâmetros fixados e $\Sigma_\beta$ é uma matriz positiva definida conhecida, de ordem $q \times q$.

(ii) $\sigma^{-2} \sim \text{Gamma}(a/2, b/2)$, em que $a > 0$ e $b > 0$ são conhecidos.

(iii) Seguindo Lin e Lee (2007), com o intuito de facilitar o procedimento de estimativa e atingir o objetivo de garantir a admissibilidade de $\phi$, visto que à medida que a ordem de um modelo AR($p$) aumenta, a abordagem sob as restrições de estacionariedade torna-se mais difícil na ausência de restrições explícitas em cada componente autoregressivo, desse modo realizamos uma reparametrização de $\phi$ como em Barndorff-Nielsen e Schou (1973):

$$\phi_k^{(k)} = \gamma_k, \quad \phi_i^{(k)} = \phi_i^{(k-1)} - \gamma_k\phi_{k-i}^{(k-1)}, \quad i = 1, 2, \ldots, k-1, \tag{3.6}$$

em que $\phi_i^{(p)}$ é o $i$-ésimo coeficiente de um processo AR($p$), sendo $\phi_i^{(p)} = \phi_i = \phi_i^{(i)} - \phi_{i+1}^{(i+1)}\phi_1^{(i)} - \phi_{i+2}^{(i+2)}\phi_2^{(i+1)} - \cdots - \phi_p^{(p)}\phi_{p-i}^{(p-1)}$, para $i = 1, \ldots, p-1$. Observe que a Equação (3.6) é uma transformação um-para-um que reparametriza $\phi = (\phi_1, \ldots, \phi_p)^{\top}$ em termos das autocorrelações parciais $\gamma = (\gamma_1, \ldots, \gamma_p)^{\top} \in \mathbb{R}^p$, em que $\mathbb{R} = [-1, 1]$. Tratamos $\gamma$ como os parâmetros reparametrizados e toda geração aleatória foi realizada em $\gamma$, depois

calculamos $\phi$ a partir dos $\gamma$ obtidos, invertendo a equação 3.6. Assim, a distribuição a priori corresponde a $\gamma \sim \mathrm{U}(-1,1)$.

(iv) Para a escolha da distribuição a priori do parâmetro $\nu$ dentro da classse de distribuições MEN, leva-se em consideração a especificidade inerente ao fator de escala de cada distribuição. Para o caso $t$ de Student existem muitas sugestões na literatura quanto a distribuição a priori mais apropriada, essas discussões são apresentadas em Geweke (1993), Fonseca, Ferreira e Migon (2008), Cabral, Lachos e Madruga (2012), Garay *et al.* (2015), entre outros.

No caso da distribuição t de Student e Slash será utilizada a sugestão dada por Congdon (2007), que consiste em $\nu \sim \mathrm{Exp}(\lambda)$ com um segundo nível de hierarquia para $\lambda$, dada por $\lambda \sim \mathrm{U}(c,d)$, em que $\mathrm{Exp}(\lambda)$ denota a distribuição exponencial com média $1/\lambda\,(\lambda > 0)$ e $\mathrm{U}(c,d)$ denota a distribuição Uniforme definida no intervalo $(c,d)$.

Consideramos para a distribuição normal contaminada como distribuição a priori $\nu \sim \mathrm{Beta}(\nu_0,\nu_1)$ e $\delta \sim \mathrm{Beta}(\delta_0,\delta_1)$, em que $\nu_0,\nu_1,\delta_0,\delta_1$ são valores positivos conhecidos.

### 3.3.2 Algoritmo MCMC

No contexto Bayesiano, os estimadores são obtidos a partir de medidas resumo da distribuição a posteriori como, por exemplo, esperança, moda, variância, etc. No entanto, devido a sua forma complexa, não é fácil aproximar estes momentos utilizando técnicas de integração numérica. Atualmente, uma forma eficiente e muito utilizada para aproximar estas integrais é por meio da geração de amostras da distribuição a posteriori, via algoritmo tipo MCMC (GAMERMAN; LOPES, 2006).

O nosso algoritmo MCMC desenvolvido para o modelo RL-AR($p$)-MEN-DI utiliza o método de supor que o vetor de variáveis latentes $VL = \{\mathbf{y}_m, \mathbf{y}_c, \mathbf{u}\}$, em que $\mathbf{u} = (u_1,\ldots,u_T)^\top$, podem ser completamente observados, dessa forma, é suficiente a obtenção da distribuição condicional completa para cada parâmetro do modelo, ou seja, determinar a distribuição de cada um dos parâmetros dado os dados observados e os parâmetros restantes, para então, extrairmos amostras dessas distribuições condicionais completas.

Seja $\mathbf{Y} = (Y_1,\ldots,Y_T)^\top$ realizações do modelo RL-AR($p$)-MEN sem censura e valores faltantes, e $\mathbf{y}_{(t-p)} = \{y_{t-1},\ldots,y_{t-p}\}$. A partir da definição da Equação (3.2), temos que a variável aleatória $\{Y_t|\mathbf{y}_{(t-p)},\theta\}$ para $t = p+1,\ldots,T$, segue uma distribuição MEN com

representação hierárquica dada por:

$$Y_t|\mathbf{y}_{(t-p)}, U_t = u_t \sim \mathrm{N}(\mu_t^*; u_t^{-1}\sigma^2),$$

$$U_t \sim H(\cdot|\nu).$$

Considerando o vetor de valores faltantes $\mathbf{y}_m$ e o vetor de valores censurados $\mathbf{y}_c$, existe então $\mathbf{y}_m + \mathbf{y}_c$ valores não observados da característica de interesse. Assim, $y_t$ é a realização da variável latente não observada $Y_t \sim$ MEN $(\mu_t^*, \sigma^2, \nu)$, $t \in \{C_m, C_c\}$. Como já mencionado, o procedimento chave de nosso algoritmo tipo MCMC consiste em considerar os dados "aumentados", acrescentando as variáveis latentes $\{\mathbf{y}_o, \mathbf{y}_m, \mathbf{y}_c, \mathbf{u}\}$, isto é, tratar o problema como se $\mathbf{y}_m$, $\mathbf{y}_c$ e $\mathbf{u}$ fossem de fato observados.

Os passos do nosso algoritmo são os seguintes:

**Passo 1.** Para cada um dos $M$ blocos de valores faltantes (B) $\mathbf{y}_m = (\mathbf{y}_{\mathrm{B}1}, \ldots, \mathbf{y}_{\mathrm{B}M})$, com $\mathbf{y}_{\mathrm{B}k} = \left\{y_{s_k+1}, \ldots, y_{s_k+m_k}\right\}$ e $\mathbf{y}_{(s_k+m_k+p)} = \left\{y_{s_k+m_k+1}, \ldots, y_{s_k+m_k+p}\right\}$ em que $m_k$ e $s_k$ representam o número de observações faltantes e sua posição no conjunto de dados, respectivamente, para o $k$-ésimo bloco de valores faltantes. Devemos gerar $\mathbf{y}_{\mathrm{B}*k}$, para $k = 1, 2, \ldots, M$, independentemente da distribuição condicional completa $\pi(\mathbf{y}_{\mathrm{B}k}|\mathbf{y}_{(s_k-p)}, \mathbf{y}_{(s_k+m_K+p)}, \mathbf{u}, \theta)$, que é obtida da seguinte forma:

- Considerando a representação estocástica da distribuição MEN, Equação (3.2), e o resultado obtido na Equação (3.5), temos:

$$\left(\mathbf{y}_{\mathrm{B}k}, \mathbf{y}_{(s_k+m_K+p)}\right) \mid \left(\mathbf{y}_{(s_k-p)}, \mathbf{u}, \theta\right) \sim \mathrm{N}_{m_k+p}(\mu_{kp}, \Sigma_{kp}).$$

Em que $\mu_{kp} = (\mu_{m_k}, \mu_p)^\top$ e $\Sigma_{kp} = \begin{pmatrix} \Sigma_{m_k \times m_k} & \Sigma_{m_k \times p} \\ \Sigma_{p \times m_k} & \Sigma_{p \times p} \end{pmatrix}$,
com $\mu_{m_k} = \left(\mu_{s_k+1}, \ldots, \mu_{s_k+m_k}\right)^\top$ e $\mu_p = \left(\mu_{s_k+m_k+1}, \ldots, \mu_{s_k+m_k+p}\right)^\top$, em que:
$\mu_{kp_{(j)}} = \mu_{s_k+j}$
$\Sigma_{kp_{(i,j)}} = \sum_{r=0}^{\min\{i-1, j-1\}} \frac{\psi_i \psi_{|i-j|+r} \sigma^2}{\mu_{s_k+\min\{i,j\}-r}}$, para $i, j = 1, \ldots, m_k + p$.

- Utilizando a decomposição condicional-marginal da distribuição normal multivariada, a distribuição de $\mathbf{y}_{\mathrm{B}k} | \left(\mathbf{y}_{(s_k-p)}, \mathbf{y}_{(s_k+m_K+p)}, \mathbf{u}, \theta\right)$ é dada por:

$$\mathrm{N}_{m_k}\left(\mu_{m_k}^*, \Sigma_{m_k}^*\right),$$

em que $\mu^*_{m_k} = \mu_{m_k} + \Sigma_{m_k \times p} \left(\Sigma_{p \times p}\right)^{-1} \left(\mathbf{y}_{(s_k+m_K+p)} - \mu_p\right)$ e

$\Sigma^*_{m_k} = \Sigma_{m_k \times m_k} - \Sigma_{m_k \times p} \left(\Sigma_{p \times p}\right)^{-1} \Sigma_{p \times m_k}$.

Portanto, nesse primeiro passo é gerado observações para os $m$ valores faltantes.

**Passo 2.** Quando $t \in C_c$, gerar observações de $Y_t$, a partir da distribuição condicional completa $\pi(y_t|\mathbf{y}_o, u_t, \beta, \phi, \sigma^2, \nu)$, que é uma distribuição normal truncada dada por

$$\mathrm{NT}(\mu^*_t, u_t^{-1}\sigma^2; \lfloor -\infty, V_t \rfloor),$$

Assim, nesse segundo passo construimos um novo vetor $\mathbf{y}^* = (y^*_1, \ldots, y^*_T)$ composto pelo vetor de valores observados $\mathbf{y}_o$ e pelas observações geradas para os casos censurados e os casos faltantes.

**Passo 3.** Para $t = p+1, \ldots, T$, gerar observações de $U_t$ da distribuição condicional completa $\pi(u_t|y^*_t, \beta, \phi, \sigma^2, \nu)$, conforme as diferentes distribuições MEN, seguindo Garay *et al.* (2015):

(a) para a distribuição Normal, definir $u_t = 1$ para $t = 1, 2, \ldots, T$;

(b) para a distribuição t de Student,

$$\mathrm{Gamma}\left(\frac{\nu+1}{2}, \frac{(y^*_t - \mu^*_t)^2}{2\sigma^2} + \frac{\nu}{2}\right);$$

(c) para a distribuição slash,

$$\mathrm{TGamma}\left(\nu + \frac{1}{2}, \frac{(y^*_t - \mu^*_t)^2}{2\sigma^2} + \frac{\nu}{2}; [0,1]\right),$$

em que TGamma representa a distribuição gamma truncada;

(d) para a distribuição normal contaminada, temos a distribuição discreta que assume os valores de $\delta$ (definido na Subseção 3.3.1) com probabilidade $pb^*$ e 1 com probabilidade $1 - pb^*$, em que

$$pb^* = \nu\delta^{(1/2)}\exp\left(-\frac{\delta}{2}\left(\frac{y^*_t - \mu^*_t}{\sigma}\right)^2\right) \quad \text{e}$$

$$1 - pb^* = (1-\nu)\exp\left(-\frac{1}{2}\left(\frac{y^*_t - \mu^*_t}{\sigma}\right)^2\right).$$

**Passo 4.** Gerar observações de $\beta$ a partir da distribuição condicional completa $\pi(\beta|y^*_t, u_t, \phi, \sigma^2, \nu)$, que é definida por

$$\mathrm{N}_q(\mathbf{D}\mathbf{A}^{-1}, \mathbf{A}^{-1}),$$

em que $\mathbf{A} = \sum_{t=p+1}^{T} \frac{U_t}{\sigma^2} \left( x_t^\top - \sum_{k=1}^{p} \phi_k x_{t-k}^\top \right)^2 + \Sigma_\beta^{-1}$

e $\mathbf{D} = \sum_{t=p+1}^{T} \frac{Z_t U_t}{\sigma^2} \left( x_t^\top - \sum_{k=1}^{p} \phi_k x_{t-k}^\top \right) + \Sigma_\beta^{-1} \mu_\beta^\top$, em que $Z_t = \left( y_t^* - \sum_{k=1}^{p} \phi_k y_{t-k}^* \right)$.

**Passo 5.** Gerar observações de $\sigma^{-2}$ a partir da distribuição condicional completa $\pi(\sigma^{-2}|y_t^*, u_t, \beta, \phi, \nu)$, que é definida por

$$\text{Gamma}\left( \frac{T+a}{2}, \frac{b + \sum_{p+1}^{T} U_t (y_t^* - \mu_t^*)^2}{2} \right).$$

**Passo 6.** Gerar observações de $\phi$, usando o método Metropolis-Hastings, a partir da distribuição condicional completa $\pi(\gamma|y_t^*, u_t, \beta, \sigma^2, \nu)$, como sugerido por Lin e Lee (2007), da seguinte forma:

(i) transformar $\gamma = (\gamma_1, \ldots, \gamma_p)^\top$ para $\gamma^* = (\gamma_1^*, \ldots, \gamma_p^*)^\top \in \mathbb{R}^p, \mathbb{R} = (-\infty, \infty)$, em que $\gamma_i^* = \log((1+\gamma_i)/(1-\gamma_i))$, com $i = 1, \ldots, p$

(ii) utilizando o método Metropolis-Hastings, gerar observações de $\gamma^*$ a partir da distribuição condicional marginal

$$\pi(\gamma^*|\mathbf{y}^*, \mathbf{u}, \boldsymbol{\beta}, \boldsymbol{\sigma^2}, \boldsymbol{\nu}) = \pi(\gamma|\mathbf{y}^*, \mathbf{u}, \boldsymbol{\beta}, \boldsymbol{\sigma^2}, \boldsymbol{\nu}) J_{\gamma^*}$$

em que $\pi(\gamma|\mathbf{y}^*, \mathbf{u}, \boldsymbol{\beta}, \boldsymbol{\sigma^2}, \boldsymbol{\nu}) = \prod_{t=p+1}^{T} f_{\text{MEN}}(y_t^*|\mu_t^*, \sigma^2, \nu)$ e $J_{\gamma^*} = \prod_{i=1}^{p} \{2\exp(\gamma_i^*)/[1+\exp(\gamma_i^*)]^2\}$ é o jacobiano da transformação de $\gamma$ para $\gamma^*$.

A distribuição proposta escolhida foi uma distribuição normal multivariada $p$-dimensional. Assim, seja uma observação $\gamma^{*(j-1)}$ obtida na fase $j-1$, gerar uma observação candidata $\gamma^{**}$ a partir da distribuição

$$\text{N}_p\left( \gamma^{*(j-1)}, r^2 \Sigma_{\gamma^*}^{(j-1)} \right)$$

em que a escala $r \approx 2{,}4/\sqrt{p}$, como sugerido por Gelman *et al.* (2004). A matriz de covariância $\Sigma_{\gamma^*}^{(j-1)}$ pode ser estimada invertendo a matriz de informação da amostra dado $\gamma^{*(j-1)}$.

A nova observação $\gamma^{**}$ é aceita com probabilidade

$$\min\left\{ \frac{\pi(\gamma^{**}|\cdots)\gamma^{**}}{\pi(\gamma^{*(j-1)}|\cdots)\gamma^{*(j-1)}}, 1 \right\}.$$

As amostras são obtidas a partir das distribuições parcialmente marginais, integrando as variáveis latentes $u_1, \ldots, u_T$.

(iii) tendo obtido $\gamma^*$ do algoritmo Metropolis-Hastings, o transformamos de volta para $\gamma$ calculando $\gamma_i = [\exp(\gamma_i^*) - 1] / [\exp(\gamma_i^*) + 1]$, com $i = 1, \ldots, p$.

(iv) transformar $\gamma$ de volta para $\phi$ invertendo a Equação (3.6).

**Passo 7.** Gerar observações de $\nu$ a partir da distribuição condicional completa $\pi(\nu|y_t^*, \beta, \phi, \sigma^2, \lambda)$.

Este procedimento de geração depende da distribuição MEN, e em alguns casos é necessário introduzir um passo Metropolis-Hastings.

Assim, por exemplo, conforme sugerido por Garay *et al.* (2015)

(a) para a distribuição *t* de Student,

(i) gerar observações de $\lambda$ a partir da distribuição condicional $\pi(\lambda|\nu)$, a qual é TGamma$(2, \nu, \lfloor c, d \rfloor)$.

(ii) utilizando o método Metropolis-Hastings, gerar observações de $\nu$ a partir da distribuição condicional marginal

$$\pi(\nu|\mathbf{y}^*, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^{\mathbf{2}}, \boldsymbol{\lambda}) \propto \exp(-\lambda\nu) \exp\left[\sum_{t=p+1}^{T} \log F_{\mathrm{TS}}\left(\frac{V_t - \mu_t^*}{\sigma}\right)^{\mathbb{I}_{\mathbf{y}_c}(y_t)} + \sum_{t=p+1}^{T} \log f_{\mathrm{TS}}\left(y_t^*|\mu_t^*, \sigma^2\right)^{1-\mathbb{I}_{\mathbf{y}_c}(y_t)}\right]. \tag{3.7}$$

As propostas são obtidas da seguinte forma: Dada uma observação $\nu^{(j-1)}$ obtida na fase $j-1$, geramos uma observação candidata $\nu^*$ da distribuição Log-normal

$$\mathrm{LN}(\log \nu^{(j-1)}, \xi_\nu^2),$$

cuja fdp é definida por:

$$q(\nu^*|\nu^{(j-1)}) = \frac{1}{\nu^* \xi_\nu \sqrt{2\pi}} \exp\left(\frac{(\log \nu^* - \nu^{(j-1)})^2}{2\xi_\nu^2}\right).$$

A nova observação $\nu^*$ é aceita com probabilidade

$$\min\left\{\frac{\pi(\nu^*|\cdots)\nu^*}{\pi(\nu^{(j-1)}|\cdots)\nu^{(j-1)}}, 1\right\}$$

em que $\pi(\nu^*|\cdots)$ representa a Equação (3.7), avaliada usando os valores atuais de $\beta, \phi, \sigma^2, \lambda$ e $y^*$. Neste caso, as amostras são obtidas a partir das distribuições parcialmente marginais, integrando as variáveis latentes $u_1, \ldots, u_T$. Este método, conhecido como princípio *colapsing*, geralmente é mais eficiente do que a amostragem da distribuição condicional completa. Para uma discussão mais detalhada veja Liu (1994) e Massuia *et al.* (2017).

(b) para a distribuição slash,

(i) gerar observações de $\lambda$ a partir da distribuição condicional $\pi(\lambda|\nu)$, a qual é TGamma$(2,\nu,\lfloor c,d\rfloor)$.

(ii) gerar observações de $\nu$ a partir da distribuição condicional $\pi(\nu|\mathbf{u})$ que é definida por

$$\text{Gamma}\left(T-p+1,\lambda-\sum_{t=p+1}^{T}\log(u_t)\right)$$

(c) para a distribuição normal contaminada,

(i) gerar uma amostra de $\nu$ a partir da distribuição condicional $\pi(\nu|\mathbf{u},\delta)$, definida como Beta$(\nu_0+z_\delta;\nu_1+T-p-z_\delta)$, em que $z_\delta=(T-p-\sum_{p+1}^{T}S_t)/(1-\delta)$ e

$$S_t=\begin{cases}1 & \text{se} \quad u_t=\delta\\ 0 & \text{se} \quad u_t\neq\delta.\end{cases}$$

(ii) considere a distribuição condicional marginal de $\delta|y^*,\beta,\phi,\sigma^2,\nu$, definida por

$$\pi\left(\delta|\mathbf{y}^*,\boldsymbol{\beta},\phi,\boldsymbol{\sigma^2},\boldsymbol{\nu}\right)\propto\delta^{\delta_0-1}(1-\delta)^{\delta_1-1}$$
$$\exp\left[\sum_{t=p+1}^{T}\log F_{\text{NC}}\left(\frac{V_t-\mu_t^*}{\sigma},\nu,\delta\right)^{\mathbb{I}_{\mathbf{y}c}(y_t)}+\sum_{t=p+1}^{T}\log f_{\text{NC}}\left(y_t^*|\mu_t^*,\sigma^2\right)^{1-\mathbb{I}_{\mathbf{y}c}(y_t)}\right],$$

em que $F_{NC}$ é a fda da normal contaminada. Dada a parametrização $\delta_r=\delta/1-\delta$, gerar observações $\delta_r$ da distribuição

$$\pi\left(\delta_r|\mathbf{y}^*,\boldsymbol{\beta},\phi,\boldsymbol{\sigma^2},\boldsymbol{\nu}\right)=\frac{1}{(1+\delta_r)^2}\pi\left(\frac{\delta_r}{1+\delta_r}\middle|\mathbf{y}^*,\boldsymbol{\beta},\phi,\boldsymbol{\sigma^2},\boldsymbol{\nu}\right),$$

utilizando o método de Metropolis-Hastings com a distribuição Log-normal, como densidade proposta. Assim, seja uma observação $\delta_r^{(j-1)}$ obtida na fase $j-1$, gerar uma observação candidata $\delta_r^*$ a partir da distribuição

$$\text{LN}(\delta_r^{(j-1)},\xi^2_{\delta_r}).$$

A nova observação $\delta^*$ é aceita com probabilidade

$$\min\left\{\frac{\pi(\delta^*|\cdots)\delta^*}{\pi(\delta^{(j-1)}|\cdots)\delta^{(j-1)}},1\right\}.$$

Neste caso, as amostras são obtidas a partir de distribuições parcialmente marginais, integrando as variáveis latentes $u_1,\ldots,u_T$.

## 3.4 CRITÉRIOS DE COMPARAÇÃO DOS MODELOS

Quando é preciso escolher entre modelos distintos para o mesmo conjunto de dados, existem várias propostas de critérios de comparação de modelos Bayesianos. Um dos critérios utilizados em trabalhos aplicados é o fator de Bayes, definido como

$$FB_{(M_i,M_j)} = \frac{f(\mathbf{y}|M_i)}{f(\mathbf{y}|M_j)},$$

em que $M_i$ e $M_j$ representam o $i$-ésimo e o $j$-ésimo modelo, respectivamente, $f(\mathbf{y}|M_i)$ e $f(\mathbf{y}|M_j)$ são as verossimilhanças marginais do $i$-ésimo e do $j$-ésimo modelo, respectivamente.

As verossimilhanças marginais são definidas como

$$f(\mathbf{y}|M_k) = \int f(\mathbf{y}|\boldsymbol{\theta}_k, M_k)\pi(\boldsymbol{\theta}_k|M_k)d\boldsymbol{\theta}_k,$$

em que $f(\mathbf{y}|\boldsymbol{\theta}_k, M_k)$ é a função de verossimilhança para o modelo $M_K$, $\pi(\boldsymbol{\theta}_k|M_k)$ a distribuição a priori e $\boldsymbol{\theta}_k$ é o vetor de parâmetros do modelo $M_K$.

Como discutido em Garay *et al.* (2020) e Lin e Lee (2007) consideraremos a densidade a posteriori, a partir da qual pode ser calculada a verossimilhança marginal, para aproximar $f(y|\theta)$, ou seja,

$$f(\mathbf{y}|M_k) \approx \left\{ \frac{1}{L} \sum_{l=1}^{L} \frac{1}{f(\mathbf{y}|\boldsymbol{\theta}_k^{(l)}, M_k)} \right\}^{-1},$$

em que $\boldsymbol{\theta}_k^{(l)}$ representa a amostra gerada da distribuição a posteriori de $\theta_k$ obtida na $l$-ésima iteração do algoritmo MCMC.

Para a comparação dos modelos, realiza-se o fator de Bayes, uma interpretação para o fator de Bayes é dada em Jeffreys (1998), em que: $FB_{(M_i,M_j)} < 1$ demonstra evidência a favor de $M_j$; $1 \leqslant FB_{(M_i,M_j)} < 3,2$ demonstra evidência muito fraca a favor de $M_i$; $3,2 \leqslant FB_{(M_i,M_j)} < 10$ demonstra evidência fraca a favor de $M_i$; $10 \leqslant FB_{(M_i,M_j)} < 100$ demonstra evidência forte a favor de $M_i$ e $FB_{(M_i,M_j)} \geqslant 100$ demonstra evidência muito forte a favor de $M_i$.

No próximo capítulo, abordaremos os resultados de três estudos de simulação, realizados em condições específicas, para avaliar e comparar os modelos propostos neste capítulo.

# 4 ESTUDOS DE SIMULAÇÃO

Elaboramos três estudos de simulação com o intuito de avaliar o desempenho das estimativas Bayesianas obtidas para os parâmetros $\beta_1, \beta_2, \phi_1, \phi_2$ e $\sigma^2$ do modelo RL-AR($p$)-MEN-DI, definido na Equação (3.2). Apesar de termos avaliado o parâmetro $\nu$, não o incluimos nos resultados tendo em vista que a comparação desse parâmetro entre todos os casos é inviável. Todos os procedimentos computacionais foram implementados utilizando o software `R` (R CORE TEAM, 2021).

Em todos os cenários de simulação geramos amostras artificiais R=500 de tamanho $T \in \{150, 300\}$ a partir do modelo RL-AR($p$)-MEN-DI com $p \in \{1, 2\}$, considerando os casos normal (N), t-Student (T), slash (SL) e normal contaminada (NC), conforme definidos na Seção (2.4), com $\mathbf{x}_t^\top = (1, x_t)$, para $t = 1, 2, \ldots, T$. Os valores $x_t$ foram fixados e obtidos a partir de amostras aleatórias da distribuição U(0, 1).

Os valores dos parâmetros foram $\boldsymbol{\beta} = (1, 4)^\top$, $\boldsymbol{\phi} = (\phi_1, \phi_2)^\top = (0{,}656, -0{,}207)^\top$ e $\sigma^2 = 1{,}5$.

Para cada réplica, foram realizadas 100.000 iterações do amostrador de Gibbs, tal que as primeiras 20.000 $(20\%)$ iterações foram descartadas como amostras *burn-in* e para eliminar potenciais problemas devido a autocorrelação, as amostras subsequentes sorteadas com espaçamento igual a 5 foram mantidas, de modo que o tamanho de cada cadeia final de observações resultou em 16.000.

A partir das 16.000 observações que formaram a cadeia final, foi calculado para cada parâmetro a média a posteriori, o desvio padrão a posteriori, o percentil $2{,}5$ a posteriori e o percentil $97{,}5$ a posteriori.

## 4.1 CENÁRIO 1

Tem como propósito verificar as consequências no comportamento das estimativas Bayesianas dos parâmetros no contexto de existência de censura no conjunto de dados. Os dados simulados seguem um modelo RL-AR($p$)-MEN-DI com $p \in \{1, 2\}$. Para realizar esse estudo, criamos dados censurados à esquerda, da seguinte forma: (i) ordenamos os dados em ordem crescente; (ii) determinamos o valor da censura pelo nível de censura escolhido para a amostra, assim, todos os valores que estavam à esquerda do valor da censura foram substituídos pelo valor da censura; (iii) retornamos os dados a sua ordem original. Foram considerados diferentes níveis de censura NCens $\in \{0\%, 5\%, 15\%, 30\%\}$ e diferentes tamanhos de amostras $T \in \{150, 300\}$.

**Tabela 1 – Resumo das estimativas Bayesianas baseadas nas 500 réplicas simuladas dos modelos RL-AR($p$)-MEN-DI, para $p = 1, T = 300$ e diferentes níveis de censura.**

| | | RL-AR(1)-N-DI | | | RL-AR(1)-T-DI | | | RL-AR(1)-SL-DI | | | RL-AR(1)-NC-DI | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **NCens** | **PR** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** |
| 0% | $\beta_1$ | 0,9947 | 0,2397 | 95,6% | 1,0194 | 0,2793 | 94,8% | 1,0077 | 0,2726 | 95,8% | 1,0076 | 0,2579 | 95,4% |
| | $\beta_2$ | 4,0105 | 0,2076 | 96,2% | 3,9898 | 0,2434 | 95,6% | 3,9978 | 0,2389 | 94,4% | 4,0033 | 0,2226 | 96,2% |
| | $\phi_1$ | 0,6561 | 0,0445 | 96,0% | 0,6529 | 0,0376 | 94,6% | 0,6504 | 0,0445 | 95,6% | 0,6526 | 0,0447 | 95,4% |
| | $\sigma^2$ | 1,5099 | 0,1249 | 93,6% | 1,4990 | 0,2057 | 94,6% | 1,5406 | 0,2311 | 99,4% | 1,3928 | 0,3784 | 100,0% |
| 5% | $\beta_1$ | 1,0078 | 0,2404 | 96,0% | 1,0134 | 0,2899 | 95,0% | 0,9945 | 0,2721 | 95,0% | 0,9852 | 0,2579 | 95,2% |
| | $\beta_2$ | 4,0145 | 0,2124 | 95,0% | 3,9988 | 0,2674 | 95,0% | 3,9971 | 0,2439 | 95,8% | 4,0057 | 0,2287 | 94,8% |
| | $\phi_1$ | 0,6499 | 0,0453 | 95,4% | 0,6513 | 0,0407 | 95,2% | 0,6432 | 0,0464 | 95,8% | 0,6491 | 0,0455 | 96,8% |
| | $\sigma^2$ | 1,5320 | 0,1333 | 93,8% | 1,5100 | 0,2162 | 95,2% | 1,5379 | 0,2440 | 99,4% | 1,4219 | 0,3550 | 99,6% |
| 15% | $\beta_1$ | 0,9651 | 0,2417 | 95,2% | 0,9950 | 0,2887 | 94,6% | 0,9887 | 0,2780 | 93,6% | 0,9848 | 0,2614 | 94,8% |
| | $\beta_2$ | 4,0257 | 0,2330 | 96,0% | 4,0014 | 0,2678 | 95,0% | 4,0307 | 0,2641 | 95,8% | 4,0114 | 0,2479 | 96,0% |
| | $\phi_1$ | 0,6315 | 0,0476 | 95,6% | 0,6386 | 0,0435 | 95,4% | 0,6318 | 0,0487 | 94,4% | 0,6330 | 0,0479 | 96,2% |
| | $\sigma^2$ | 1,5836 | 0,1529 | 92,2% | 1,5689 | 0,2389 | 95,0% | 1,6163 | 0,2604 | 98,8% | 1,5162 | 0,3421 | 100,0% |
| 30% | $\beta_1$ | 0,9888 | 0,2587 | 93,0% | 0,9843 | 0,3151 | 94,4% | 0,9804 | 0,2938 | 95,0% | 0,9962 | 0,2786 | 94,6% |
| | $\beta_2$ | 4,0263 | 0,2727 | 96,2% | 4,0112 | 0,3144 | 94,4% | 4,0394 | 0,3107 | 97,2% | 4,0268 | 0,2936 | 96,0% |
| | $\phi_1$ | 0,6062 | 0,0517 | 88,0% | 0,6201 | 0,0475 | 92,4% | 0,6046 | 0,0528 | 89,4% | 0,6067 | 0,0521 | 91,4% |
| | $\sigma^2$ | 1,6884 | 0,1935 | 84,2% | 1,6795 | 0,2831 | 94,6% | 1,7127 | 0,2942 | 97,6% | 1,6466 | 0,3595 | 99,6% |

Fonte: Elaborada pelo autor (2022).
Simulação 01. Nível de Censura (NCens), Parâmetro (PR), Média das médias das 500 réplicas (M Med), Média dos desvios padrão das 500 réplicas (M DP), Percentual de cobertura do intervalo de credibilidade de 95% (P Cob).

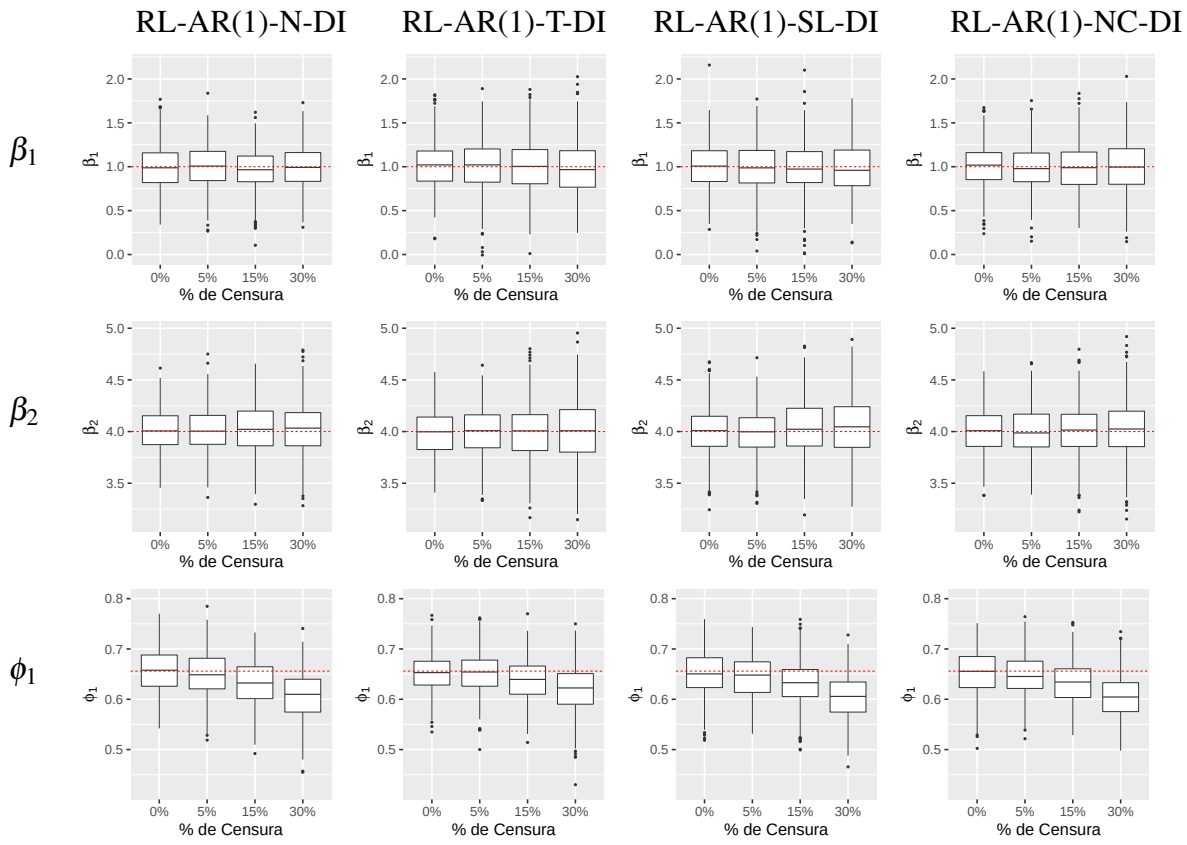
**Tabela 2 – Resumo das estimativas Bayesianas baseadas nas 500 réplicas simuladas dos modelos RL-AR($p$)-MEN-DI, para $p = 2, T = 300$ e diferentes níveis de censura.**

| | | RL-AR(2)-N-DI | | | RL-AR(2)-T-DI | | | RL-AR(2)-Sl-DI | | | RL-AR(2)-NC-DI | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **NCens** | **PR** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** |
| 0% | $\beta_1$ | 1,0005 | 0,1631 | 95,6% | 0,9943 | 0,1897 | 95,2% | 1,0139 | 0,1866 | 94,8% | 0,9972 | 0,1758 | 95,0% |
| | $\beta_2$ | 4,0041 | 0,1948 | 94,8% | 3,9980 | 0,2280 | 95,6% | 3,9873 | 0,2229 | 94,6% | 4,0050 | 0,2105 | 96,8% |
| | $\phi_1$ | 0,6522 | 0,0572 | 94,2% | 0,6573 | 0,0477 | 93,8% | 0,6557 | 0,0570 | 94,8% | 0,6523 | 0,0572 | 95,8% |
| | $\phi_2$ | -0,2030 | 0,0571 | 93,4% | -0,2098 | 0,0476 | 94,0% | -0,2065 | 0,0569 | 95,2% | -0,2050 | 0,0573 | 95,8% |
| | $\sigma^2$ | 1,5190 | 0,1261 | 93,2% | 1,4683 | 0,2021 | 93,0% | 1,5320 | 0,2310 | 99,6% | 1,3908 | 0,3785 | 99,4% |
| 5% | $\beta_1$ | 0,9913 | 0,1666 | 96,2% | 0,9920 | 0,1944 | 95,6% | 1,0090 | 0,1902 | 95,0% | 0,9913 | 0,1792 | 95,8% |
| | $\beta_2$ | 4,0168 | 0,2019 | 95,6% | 4,0041 | 0,2366 | 96,4% | 3,9973 | 0,2308 | 95,0% | 4,0139 | 0,2179 | 96,2% |
| | $\phi_1$ | 0,6425 | 0,0585 | 94,6% | 0,6397 | 0,0532 | 95,0% | 0,6422 | 0,0597 | 96,0% | 0,6417 | 0,0588 | 94,8% |
| | $\phi_2$ | -0,1967 | 0,0586 | 94,0% | -0,1952 | 0,0514 | 94,2% | -0,1972 | 0,0592 | 94,8% | -0,1977 | 0,0588 | 96,4% |
| | $\sigma^2$ | 1,5466 | 0,1352 | 93,2% | 1,5050 | 0,2163 | 93,2% | 1,5489 | 0,2444 | 99,4% | 1,4437 | 0,3564 | 99,6% |
| 15% | $\beta_1$ | 0,9946 | 0,1764 | 96,0% | 0,9895 | 0,2056 | 96,0% | 1,0039 | 0,2006 | 96,0% | 0,9856 | 0,1898 | 95,6% |
| | $\beta_2$ | 4,0117 | 0,2218 | 95,2% | 4,0106 | 0,2561 | 97,4% | 4,0039 | 0,2522 | 95,4% | 4,0220 | 0,2386 | 96,0% |
| | $\phi_1$ | 0,6126 | 0,0623 | 92,0% | 0,6156 | 0,0576 | 92,2% | 0,6137 | 0,0638 | 92,2% | 0,6134 | 0,0626 | 91,4% |
| | $\phi_2$ | -0,1768 | 0,0623 | 94,4% | -0,1754 | 0,0544 | 92,6% | -0,1784 | 0,0630 | 95,0% | -0,1782 | 0,0624 | 95,8% |
| | $\sigma^2$ | 1,5941 | 0,1546 | 91,4% | 1,5888 | 0,2413 | 94,2% | 1,6129 | 0,2609 | 99,0% | 1,5318 | 0,3416 | 99,8% |
| 30% | $\beta_1$ | 0,9788 | 0,2043 | 96,2% | 0,9708 | 0,2369 | 96,0% | 0,9879 | 0,2300 | 96,2% | 0,9732 | 0,2188 | 96,8% |
| | $\beta_2$ | 4,0312 | 0,2662 | 96,0% | 4,0313 | 0,3036 | 96,4% | 4,0230 | 0,3003 | 95,8% | 4,0415 | 0,2851 | 97,4% |
| | $\phi_1$ | 0,5628 | 0,0684 | 77,2% | 0,5741 | 0,0620 | 78,4% | 0,5638 | 0,0691 | 83,6% | 0,5630 | 0,0686 | 80,2% |
| | $\phi_2$ | -0,1458 | 0,0684 | 90,6% | -0,1456 | 0,0577 | 87,8% | -0,1467 | 0,0683 | 93,8% | -0,1447 | 0,0681 | 90,2% |
| | $\sigma^2$ | 1,7021 | 0,1947 | 83,2% | 1,7261 | 0,2905 | 89,0% | 1,7359 | 0,3011 | 98,2% | 1,6593 | 0,3609 | 100% |

Fonte: Elaborada pelo autor (2022).
Simulação 01. Nível de Censura (NCens), Parâmetro (PR), Média das médias das 500 réplicas (M Med), Média dos desvios padrão das 500 réplicas (M DP), Percentual de cobertura do intervalo de credibilidade de 95% (P Cob).

Comparando os resultados das estimativas Bayesianas dos modelos RL-AR($p$)-MEN-DI, com $p \in \{1,2\}$, a partir das Tabelas 1 e 2 verificamos que o algoritmo MCMC proposto produz estimativas aproximadas ao valor verdadeiro em quase todas as proporções de censura e que há uma sensibilidade nas variações das estimativas com o aumento da censura, neste sentido, o parâmetro $\phi_1$ apresenta-se com maior sensibilidade ao percentual de censura de 30%, seguido por $\sigma^2$ que também aponta certa sensibilidade.

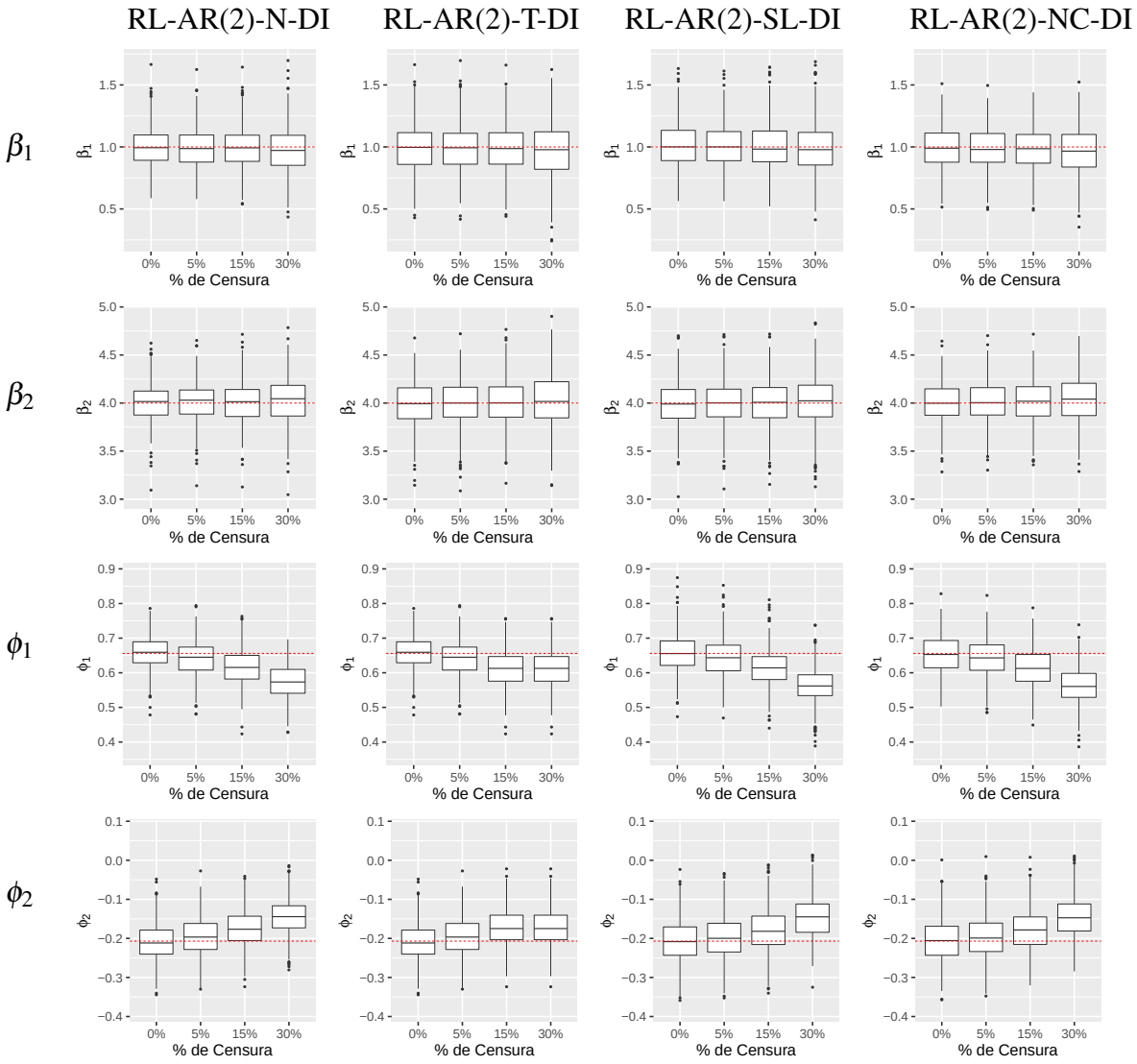
**Figura 2 – Estimativas Bayesianas médias de $\beta_1$, $\beta_2$ e $\phi_1$ para os modelos RL-AR(1)-MEN-DI, com $T = 300$ e diferente níveis de censura.**




Fonte: Elaborada pelo autor (2022).

Nas Figuras 2 e 3, cuja linha vermelha representa o valor verdadeiro dos parâmetros, podemos observar como estão distribuídos as estimativas Bayesianas dos parâmetros em cada modelo, assim como aqueles que mais se afastam do valor verdadeiro, identificando um maior afastamento quando o nível de censura é de 30%. Os resultados para o caso $T = 150$ são semelhantes e são apresentados no apêndice A. Contudo, é possível perceber que com o aumento da amostra o descréscimo do percentual de cobertura dos intervalos de credibilidade a posteriori de 95% torna-se evidente com uma censura ao nível de 30%, uma vez que ao aumentar a amostra, também aumenta-se a quantidade de valores censurados.

**Figura 3 – Estimativas Bayesianas médias de $\beta_1$, $\beta_2$, $\phi_1$ e $\phi_2$ para os modelos RL-AR(2)-MEN-DI, com $T = 300$ e diferente níveis de censura.**




Fonte: Elaborada pelo autor (2022).

## 4.2 CENÁRIO 2

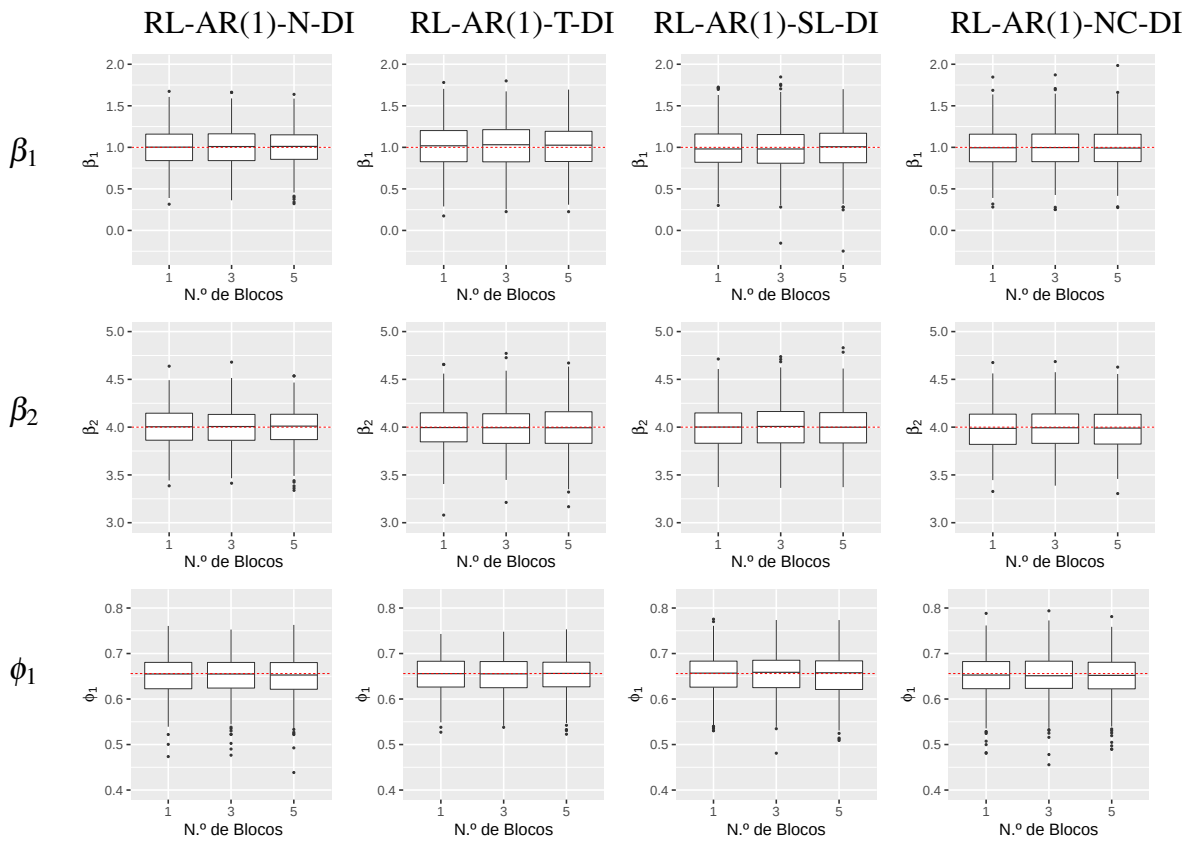
O objetivo deste cenário é comparar o desempenho das estimativas Bayesianas dos parâmetros para o modelo RL-AR($p$)-MEN-DI com $p \in \{1,2\}$ na presença de *missing values* (*NA*) na variável resposta. Neste cenário, construimos blocos de *NAs* (B) em posições diferentes para cada réplica, para isso (i) sorteamos valores aleatórios da amostra de forma que cada valor sorteado corresponderia ao primeiro dado ausente do bloco de NAs, formado por 5 valores ausentes consecutivos por bloco; (ii) quando B $> 1$, os blocos construídos, assim como os seus $p$ valores anteriores e os seus $p$ valores posteriores, não participavam do próximo sorteio. Então, assumimos diferentes quantidades de blocos de *NAs* B $\in \{1,3,5\}$ e o tamanho de amostra

$T = 300$, que representam 1,6%, 5% e 8% da amostra.

As Tabelas 3 e 4 apresentam as estimativas Bayesianas dos parâmetros para os modelos RL-AR($p$)-MEN-DI, com $p \in \{1,2\}$. Observamos que há certa similaridade nos resultados das médias das estimativas Bayesianas apesar da distinção no número de blocos de dados ausentes, bem como para o percentual de cobertura dos intervalos de credibilidade de 95%, nos quais os resultados também são semelhantes, com as coberturas em torno 95%.

Os boxplots, apresentados nas Figuras 4 e 5, indicam que, em geral, o nosso algoritmo proposto produziu as estimativas bayesinas mais próximas dos valores verdadeiros para todos os parâmetros, demonstrando, assim, evidências de que o algoritmo produz estimativas precisas mesmo na presença de valores ausentes.

**Figura 4 – Estimativas Bayesianas médias de $\beta_1$, $\beta_2$ e $\phi_1$ para os modelos RL-AR(1)-MEN-DI, com $T = 300$ e diferente número de blocos de valores ausentes.**




Fonte: Elaborada pelo autor (2022).

**Tabela 3 – Resumo das estimativas Bayesianas baseadas nas 500 réplicas simuladas dos modelos RL-AR($p$)-MEN-DI, para $p = 1, T = 300$ e diferentes número de blocos de valores ausentes.**

| | | RL-AR(1)-N-DI | | | RL-AR(1)-T-DI | | | RL-AR(1)-SL-DI | | | RL-AR(1)-NC-DI | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **N. ° de Blocos** | **PR** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** |
| | $\beta_1$ | 1,0009 | 0,2352 | 94,6% | 1,0140 | 0,2763 | 95,4% | 0,9859 | 0,2703 | 95,0% | 1,0001 | 0,2542 | 96,2% |
| 1 | $\beta_2$ | 4,0037 | 0,2012 | 96,0% | 3,9974 | 0,2364 | 94,4% | 3,9902 | 0,2303 | 94,0% | 3,9858 | 0,2173 | 96,0% |
| (1,6%) | $\phi_1$ | 0,6512 | 0,0451 | 95,8% | 0,6534 | 0,0379 | 95,4% | 0,6528 | 0,0447 | 95,0% | 0,6509 | 0,0451 | 94,8% |
| | $\sigma^2$ | 1,5179 | 0,1267 | 95,2% | 1,4876 | 0,2056 | 95,6% | 1,5238 | 0,2316 | 98,0% | 1,3851 | 0,3717 | 99,6% |
| | $\beta_1$ | 1,0027 | 0,2372 | 95,2% | 1,0168 | 0,2791 | 95,0% | 0,9784 | 0,2742 | 95,2% | 1,0011 | 0,2562 | 96,4% |
| 3 | $\beta_2$ | 4,0018 | 0,2051 | 96,0% | 3,9941 | 0,2410 | 95,4% | 4,0020 | 0,2346 | 95,8% | 3,9894 | 0,2214 | 95,8% |
| (5%) | $\phi_1$ | 0,6510 | 0,0458 | 95,0% | 0,6531 | 0,0387 | 94,4% | 0,6545 | 0,0454 | 95,6% | 0,6505 | 0,0457 | 94,4% |
| | $\sigma^2$ | 1,5180 | 0,1290 | 95,2% | 1,4872 | 0,2092 | 94,8% | 1,5232 | 0,2355 | 97,6% | 1,3760 | 0,3776 | 99,6% |
| | $\beta_1$ | 1,0021 | 0,2390 | 95,4% | 1,0150 | 0,2830 | 95,6% | 0,9874 | 0,2754 | 94,0% | 1,0038 | 0,2585 | 95,6% |
| 5 | $\beta_2$ | 4,0023 | 0,2096 | 96,4% | 3,9918 | 0,2465 | 94,0% | 3,9993 | 0,2396 | 95,6% | 3,9847 | 0,2261 | 94,8% |
| (8%) | $\phi_1$ | 0,6496 | 0,0466 | 95,0% | 0,6531 | 0,0393 | 95,4% | 0,6521 | 0,0462 | 95,8% | 0,6499 | 0,0466 | 94,2% |
| | $\sigma^2$ | 1,5189 | 0,1314 | 96,0% | 1,4908 | 0,2139 | 94,8% | 1,5261 | 0,2383 | 98,2% | 1,3763 | 0,3825 | 99,4% |

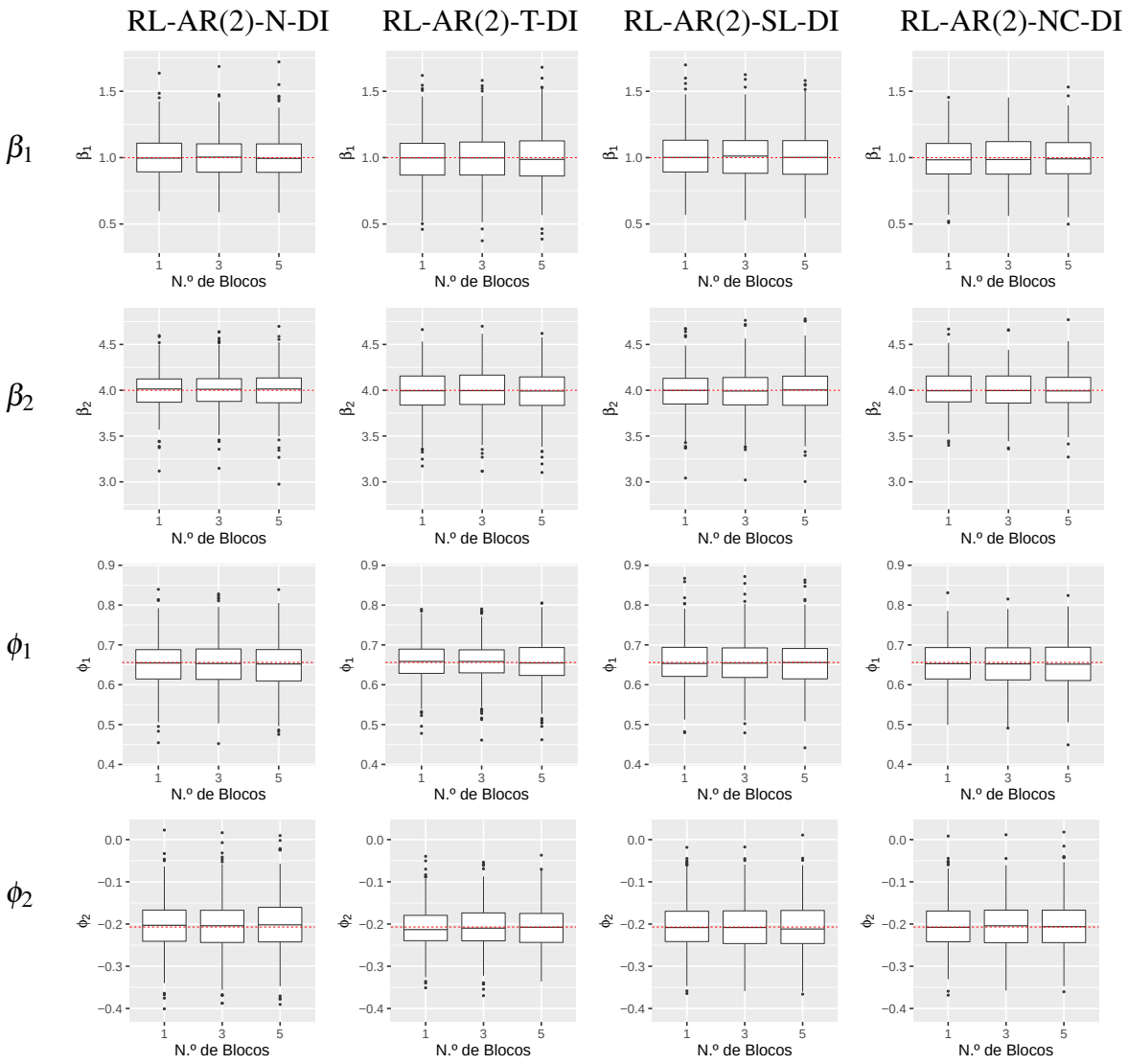
Fonte: Elaborada pelo autor (2022).
Simulação 02. Nível de Censura (NC), Parâmetro (PR), Média das médias das 500 réplicas (M Med), Média dos desvios padrão das 500 réplicas (M DP), Percentual de cobertura do intervalo de credibilidade de 95% (P Cob).

**Tabela 4 – Resumo das estimativas Bayesianas baseadas nas 500 réplicas simuladas dos modelos RL-AR($p$)-MEN-DI, para $p = 2$, $T = 300$ e diferentes número de blocos de valores ausentes.**

| | | RL-AR(2)-N-DI | | | RL-AR(2)-T-DI | | | RL-AR(2)-SL-DI | | | RL-AR(2)-NC-DI | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **N. ° de Blocos** | **PR** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** |
| 1 (1,6%) | $\beta_1$ | 1,0018 | 0,1643 | 95,4% | 0,9941 | 0,1916 | 96,0% | 1,0155 | 0,1880 | 95,4% | 0,9948 | 0,1773 | 95,4% |
| | $\beta_2$ | 4,0022 | 0,1968 | 95,4% | 3,9991 | 0,2304 | 96,0% | 3,9858 | 0,2250 | 94,2% | 4,0071 | 0,2125 | 96,2% |
| | $\phi_1$ | 0,6512 | 0,0578 | 94,8% | 0,6576 | 0,0483 | 93,4% | 0,6550 | 0,0575 | 95,8% | 0,6519 | 0,0577 | 95,0% |
| | $\phi_2$ | -0,2029 | 0,0580 | 93,8% | -0,2096 | 0,0484 | 94,6% | -0,2065 | 0,0576 | 94,0% | -0,2047 | 0,0578 | 95,8% |
| | $\sigma^2$ | 1,5205 | 0,1274 | 94,4% | 1,4726 | 0,2045 | 92,8% | 1,5334 | 0,2324 | 99,6% | 1,3877 | 0,3803 | 99,4% |
| 3 (5%) | $\beta_1$ | 1,0010 | 0,1666 | 95,4% | 0,9924 | 0,1947 | 94,4% | 1,0138 | 0,1914 | 95,6% | 0,9956 | 0,1801 | 95,6% |
| | $\beta_2$ | 4,0064 | 0,2005 | 94,8% | 3,9998 | 0,2351 | 95,4% | 3,9880 | 0,2296 | 93,4% | 4,0048 | 0,2169 | 97,2% |
| | $\phi_1$ | 0,6524 | 0,0590 | 94,2% | 0,6571 | 0,0494 | 93,4% | 0,6548 | 0,0586 | 96,6% | 0,6513 | 0,0589 | 94,4% |
| | $\phi_2$ | -0,2049 | 0,0592 | 94,0% | -0,2095 | 0,0496 | 95,6% | -0,2056 | 0,0590 | 95,2% | -0,2043 | 0,0593 | 95,4% |
| | $\sigma^2$ | 1,5084 | 0,1293 | 94,6% | 1,4698 | 0,2081 | 92,0% | 1,5309 | 0,2358 | 99,2% | 1,3880 | 0,3788 | 99,4% |
| 5 (8%) | $\beta_1$ | 0,9998 | 0,1698 | 95,8% | 0,9947 | 0,1976 | 95,2% | 1,0069 | 0,1946 | 95,0% | 0,9969 | 0,1834 | 96,4% |
| | $\beta_2$ | 4,0062 | 0,2050 | 94,8% | 3,9917 | 0,2393 | 95,0% | 3,9941 | 0,2345 | 95,4% | 4,0011 | 0,2216 | 96,4% |
| | $\phi_1$ | 0,6510 | 0,0603 | 94,8% | 0,6556 | 0,0505 | 91,6% | 0,6545 | 0,0598 | 94,6% | 0,6509 | 0,0602 | 94,8% |
| | $\phi_2$ | -0,2021 | 0,0609 | 94,6% | -0,2089 | 0,0508 | 94,4% | -0,2069 | 0,0605 | 96,2% | -0,2044 | 0,0607 | 95,6% |
| | $\sigma^2$ | 1,5150 | 0,1316 | 94,0% | 1,4609 | 0,2108 | 92,8% | 1,5288 | 0,2389 | 99,4% | 1,3892 | 0,3821 | 99,6% |

Fonte: Elaborada pelo autor (2022).
Simulação 02. Nível de Censura (NC), Parâmetro (PR), Média das médias das 500 réplicas (M Med), Média dos desvios padrão das 500 réplicas (M DP), Percentual de cobertura do intervalo de credibilidade de 95% (P Cob).

**Figura 5 – Estimativas Bayesianas médias de $\beta_1$, $\beta_2$, $\phi_1$ e $\phi_2$ para os modelos RL-AR(2)-MEN-DI, com $T = 300$ e diferente número de blocos de valores ausentes.**




Fonte: Elaborada pelo autor (2022).

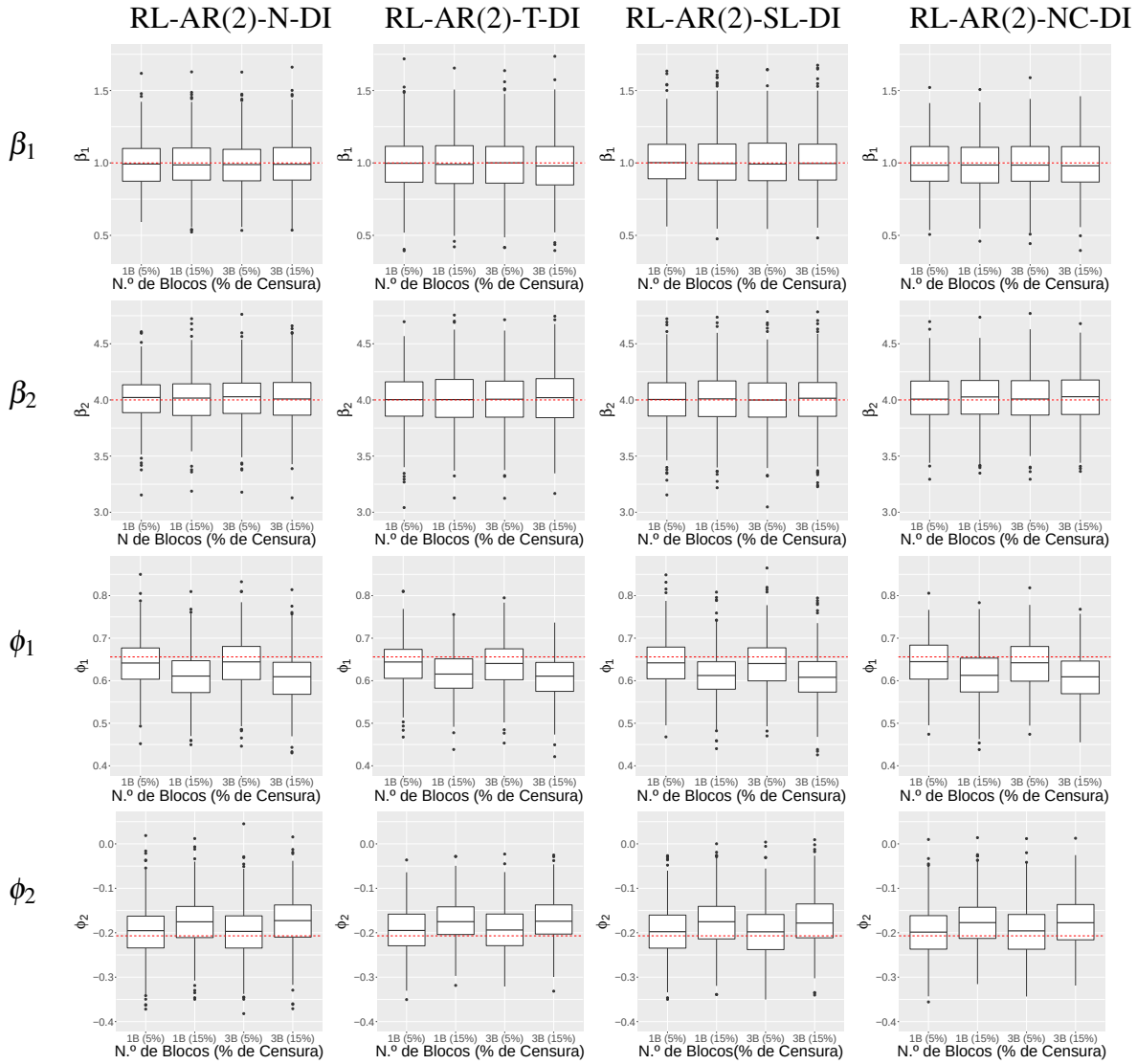
## 4.3 CENÁRIO 3

O intuito do cenário 3 é avaliar as estimativas Bayesianas dos parâmetros do modelo RL-AR(2)-MEN-DI na presença de censura e *missing values* (*NA*) na variável resposta, simultaneamente. Os blocos de valores ausentes foram produzidos utilizando método semelhante ao informado no cenário 2, neste caso, envolvendo os dados censurados, de forma a garantir que a posição dos valores ausentes não coincidissem com a posição dos valores censurados. Consideramos diferentes níveis de censura NCens $\in \{5\%, 15\%\}$, diferentes quantidades de blocos B $\in \{1, 3\}$ e o tamanho de amostra $T = 300$.

O resultado demonstrado na Tabela 5 e na Figura 6 mostram que o algoritmo funciona

muito bem diante de dados censurados e valores ausentes, apresentando resultados semelhantes nas estimativas Bayesianas e no percentual de cobertura dos intervalos de credibilidade de 95% em todos os níveis de censura e número de blocos com valores ausentes.

**Figura 6 – Estimativas Bayesianas médias de $\beta_1$, $\beta_2$, $\phi_1$ e $\phi_2$ para os modelos RL-AR(2)-MEN-DI, com $T = 300$, diferente níveis de censura e diferente número de blocos de valores ausentes.**




Fonte: Elaborada pelo autor (2022).

A partir dos três cenários avaliados, percebemos que embora os resultados considerando o nosso algoritmo proposto sejam satisfatórios, quando o nível de censura é alto (30%), os resultados perdem a precisão, em especial nos parâmetros $\sigma^2$ e $\phi_1$.

Além disso, constatamos que não há, na literatura acadêmica, estudos com modelos de regressão linear com erros autoregressivos com inovações seguindo classes de distribuições de cauda pesada, considerando dados censurados e/ou dados ausentes.

**Tabela 5 – Resumo das estimativas Bayesianas baseadas nas 500 réplicas simuladas dos modelos RL-AR($p$)-MEN-DI, para $p = 2, T = 300$, diferente níveis de censura e diferente número de blocos de valores ausentes.**

| | | | RL-AR(2)-N-DI | | | RL-AR(2)-T-DI | | | RL-AR(2)-SL-DI | | | RL-AR(2)-NC-DI | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **N. ° de Blocos** | **NCens** | **PR** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** |
| 1 (1,6%) | 5% | $\beta_1$ | 0,9926 | 0,1678 | 96,2% | 0,9940 | 0,1962 | 95,6% | 1,0078 | 0,1919 | 96,0% | 0,9906 | 0,1808 | 95,0% |
| | | $\beta_2$ | 4,0154 | 0,2041 | 95,8% | 4,0018 | 0,2392 | 95,4% | 4,0009 | 0,2332 | 95,0% | 4,0145 | 0,2201 | 96,0% |
| | | $\phi_1$ | 0,6412 | 0,0592 | 94,2% | 0,6391 | 0,0540 | 95,0% | 0,6418 | 0,0604 | 95,6% | 0,6411 | 0,0592 | 95,4% |
| | | $\phi_2$ | -0,1962 | 0,0593 | 93,6% | -0,1947 | 0,0518 | 95,0% | -0,1969 | 0,0600 | 94,8% | -0,1970 | 0,0593 | 95,6% |
| | | $\sigma^2$ | 1,5476 | 0,1366 | 94,2% | 1,5054 | 0,2184 | 93,6% | 1,5500 | 0,2463 | 99,6% | 1,4458 | 0,3556 | 99,8% |
| | 15% | $\beta_1$ | 0,9948 | 0,1783 | 95,6% | 0,9880 | 0,2076 | 95,4% | 1,0049 | 0,2028 | 96,2% | 0,9847 | 0,1917 | 96,0% |
| | | $\beta_2$ | 4,0130 | 0,2248 | 95,8% | 4,0121 | 0,2589 | 97,0% | 4,0040 | 0,2553 | 95,2% | 4,0236 | 0,2414 | 96,8% |
| | | $\phi_1$ | 0,6106 | 0,0630 | 92,0% | 0,6150 | 0,0578 | 92,4% | 0,6120 | 0,0644 | 91,8% | 0,6118 | 0,0633 | 91,6% |
| | | $\phi_2$ | -0,1756 | 0,0630 | 94,2% | -0,1740 | 0,0546 | 91,4% | -0,1765 | 0,0637 | 93,6% | -0,1769 | 0,0631 | 96,0% |
| | | $\sigma^2$ | 1,5966 | 0,1567 | 92,0% | 1,5856 | 0,2434 | 93,2% | 1,6134 | 0,2627 | 99,6% | 1,5334 | 0,3420 | 99,8% |
| 3 (5%) | 5% | $\beta_1$ | 0,9906 | 0,1708 | 95,6% | 0,9933 | 0,1996 | 94,8% | 1,0082 | 0,1977 | 96,2% | 0,9901 | 0,1840 | 96,0% |
| | | $\beta_2$ | 4,0180 | 0,2081 | 94,0% | 4,0048 | 0,2445 | 96,2% | 3,9982 | 0,2384 | 96,0% | 4,0161 | 0,2248 | 96,6% |
| | | $\phi_1$ | 0,6424 | 0,0603 | 94,2% | 0,6371 | 0,0548 | 94,2% | 0,6397 | 0,0616 | 95,6% | 0,6396 | 0,0606 | 95,0% |
| | | $\phi_2$ | -0,1969 | 0,0606 | 94,4% | -0,1933 | 0,0531 | 94,4% | -0,1958 | 0,0615 | 95,0% | -0,1956 | 0,0609 | 96,2% |
| | | $\sigma^2$ | 1,5478 | 0,1392 | 93,6% | 1,5039 | 0,2228 | 94,0% | 1,5509 | 0,2496 | 99,4% | 1,4424 | 0,3581 | 99,6% |
| | 15% | $\beta_1$ | 0,9950 | 0,1818 | 96,0% | 0,9842 | 0,2116 | 95,8% | 1,0057 | 0,2066 | 96,0% | 0,9888 | 0,1956 | 95,6% |
| | | $\beta_2$ | 4,0164 | 0,2304 | 95,4% | 4,0184 | 0,2653 | 97,2% | 4,0068 | 0,2617 | 95,8% | 4,0245 | 0,2475 | 96,8% |
| | | $\phi_1$ | 0,6069 | 0,0642 | 90,2% | 0,6094 | 0,0593 | 92,0% | 0,6078 | 0,0655 | 91,4% | 0,6083 | 0,0646 | 90,4% |
| | | $\phi_2$ | -0,1734 | 0,0644 | 93,6% | -0,1711 | 0,0559 | 92,4% | -0,1740 | 0,0650 | 94,0% | -0,1745 | 0,0646 | 94,6% |
| | | $\sigma^2$ | 1,6015 | 0,1604 | 91,2% | 1,5868 | 0,2484 | 94,8% | 1,6154 | 0,2671 | 99,4% | 1,5359 | 0,3451 | 99,8% |

Fonte: Elaborada pelo autor (2022).
Simulação 03. Nível de Censura (NCens), Parâmetro (PR), Média das médias das 500 réplicas (M Med), Média dos desvios padrão das 500 réplicas (M DP), Percentual de cobertura do intervalo de credibilidade de 95% (P Cob).

# 5 APLICAÇÃO A DADOS REAIS

Para avaliar o algoritmo MCMC proposto na seção (3.3) e estimar os parâmetros do modelo RL-AR-AR($p$)-MEN-DI utilizamos o conjunto de dados relativo à concentração total de fósforo na água do rio Iowa, EUA, analisado previamente nos estudos de Wang e Chan (2018) e Schumacher, Lachos e Dey (2017), desde um ponto de vista frequentista.

Existe hoje por parte dos ambientalistas e da população em geral, uma maior conscientização no que se refere a preservação e qualidade da água, o líquido mais precioso do planeta e componente vital no sistema de sustentação da vida. Alguns fatores, como a poluição, contribuem para o aumento de nutrientes na água. O fósforo excessivo na água do rio pode resultar em eutrofização, a multiplicação demasiada de algas, resultando em diversos problemas ambientais, sendo, portanto, um dos nutrientes de maior preocupação na água do rio Iowa como sugerido por Wang e Chan (2018).

Uma das causas do aumento da quantidade de fósforo total (P) na água também está relacionada com a vazão de água (Q), obtida do site do Serviço Geológico dos EUA. Como apontado por Libra, Wolter e Langel (2004) a concentração de fósforo na água do rio tem sido monitorada de perto no âmbito do programa de qualidade da água, conduzido pelo Departamento de Recursos Naturais de Iowa.

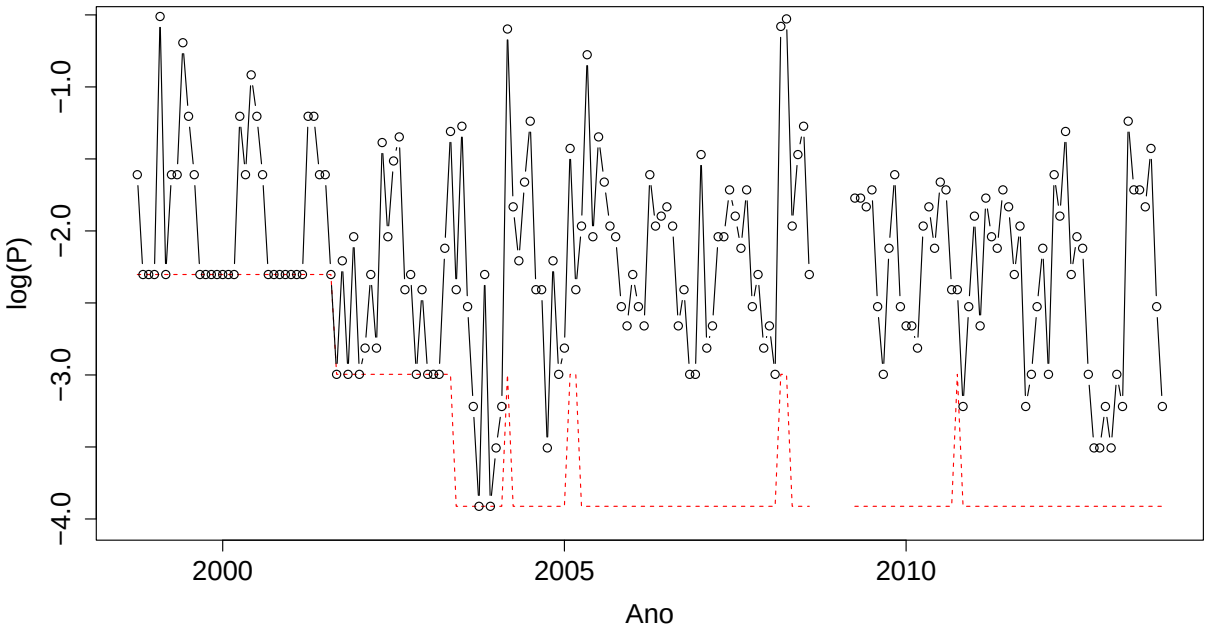
A série de concentrações mensais de fósforo (P) em mg/L, foram coletadas no período de outubro 1998 a outubro de 2013, nas águas do rio Iowa, localizado em Whitebreast Creek perto de Knoxville, Iowa, EUA. Os dados apresentados na Figura 7 contêm $15,5\%$ de observações censuradas à esquerda, representadas pela linha vermelha, ou seja, quando log(P) encontra-se abaixo de certos limites de detecção que variam ao longo do tempo e 7 (3,9%) observações faltantes, retratada pela lacuna de dados devido a interrupção do programa por falta de financiamento.

Seguindo Wang e Chan (2018), assim como Schumacher, Lachos e Dey (2017) consideramos a transformação logarítmica para P e Q, e as interações com o log da vazão de água ($\mathrm{Q}_t$), resultando no seguinte modelo:

$$\log(\mathrm{P}_t) = \sum_{j=1}^{4} [\beta_{1,j} S_{j,t} + \beta_{2,j} S_{j,t} \log(\mathrm{Q}_t)] + \varepsilon_t$$

em que $S_j$ é a variável indicadora de trimestre, com $j = 1,2,3$ e 4, tal que $S_{j,t} = 1$ se a $t$-ésima observação pertencer ao $j$-ésimo trimestre e $S_{j,t} = 0$ caso contrário, com o primeiro trimestre compreendendo janeiro a março, o segundo trimestre de abril a junho, etc; e $\varepsilon_t$ é a inovação do

**Figura 7 – Série temporal do logaritmo da concentração de fósforo (linha preta) e limites de detecção da censura (linha vermelha tracejada).**




Fonte: Elaborada pelo autor (2022).

modelo.

É importante ressaltar que nos estudos de Wang e Chan (2018) e Schumacher, Lachos e Dey (2017) $\varepsilon_t$ foi considerado como uma inovação normal de um modelo de regressão linear autoregressivo, ambos trabalharam com dados censurados e dados faltantes. No caso dos dados faltantes, Wang e Chan (2018) considerou as respostas ausentes como resultantes da censura à esquerda com um limite de censura infinito, Schumacher, Lachos e Dey (2017), considerou as respostas faltantes como resultantes de censura intervalar $(-\infty, \infty)$.

Neste trabalho consideramos $\varepsilon_t$ como uma inovação seguindo uma distribuição mistura de escala da normal definido na Equação (3.1) e substituimos o bloco de valores ausentes por meio de uma distribuição conjunta condicional como descrito na Seção (3.3). Consideramos todos os casos, apresentados na Seção (2.4), para o modelo RL-AR($p$)-MEN-DI, com $p \in \{1,2,3\}$ para análise comparativa e preditiva. Para cada modelo, geramos $L = 150.000$ iterações do amostrador de Gibbs, descartamos as primeiras 30.000 iterações como amostras de *burn-in* e escolhemos a amostra final com um espaçamento de tamanho 5, com o intuito de reduzir a autocorrelação, resultando em 24.000 realizações finais.

O critério do fator de Bayes é uma ferramenta para comparar modelos em diversos contextos, como descrito na Seção (3.4), é um recurso que envolve o cálculo das densidades marginais $f(y|\theta_k)$ para cada modelo $(M_K)$. Nesta aplicação, obtemos para cada caso do modelo

MEN-AR($p$) as $f(y|M_k)$, os resultados estão apresentados na tabela 6.

**Tabela 6 – Estimativa das densidades marginais para os modelos RL-AR($p$)-MEN-DI.**

| Modelo ($M_k$) | $f(y\|M_k)$ |
|---|---|
| RL-AR(1)-N-DI | $8,1636 \times e^{-63}$ |
| RL-AR(2)-N-DI | $3,1168 \times e^{-63}$ |
| RL-AR(3)-N-DI | $9,0565 \times e^{-64}$ |
| RL-AR(1)-T-DI | $7,3787 \times e^{-64}$ |
| RL-AR(2)-T-DI | $1,5810 \times e^{-62}$ |
| RL-AR(3)-T-DI | $1,2164 \times e^{-61}$ |
| RL-AR(1)-SL-DI | $3,4479 \times e^{-64}$ |
| RL-AR(2)-SL-DI | $5,6672 \times e^{-62}$ |
| RL-AR(3)-SL-DI | $2,1892 \times e^{-62}$ |
| RL-AR(1)-NC-DI | $2,8059 \times e^{-71}$ |
| RL-AR(2)-NC-DI | $3,0997 \times e^{-63}$ |
| RL-AR(3)-NC-DI | $4,4727 \times e^{-69}$ |

Fonte: Elaborada pelo autor (2022).

Selecionando o modelo RL-AR(3)-SL-DI para realizar o fator de Bayes $(FB_{(M_i,M_j)})$ com os demais modelos, obtivemos os seguintes resultados em relação aos modelos RL-AR(1)-N-DI, RL-AR(2)-N-DI, RL-AR(3)-N-DI, RL-AR(1)-T-DI, RL-AR(2)-T-DI, RL-AR(3)-T-DI, RL-AR(1)-SL-DI, RL-AR(2)-SL-DI, RL-AR(1)-NC-DI, RL-AR(2)-NC-DI e RL-AR(3)-NC-DI: 2,681623, 7,023905, 24,1724, 29,66908, 1,384647, 0,17997, 63,49285, 0,3862925, $7,8019 \times 10^8$, 7,062574 e $4,894577 \times 10^6$, respectivamente. Observamos evidências a favor do modelo RL-AR(3)-SL-DI, aparentemente mais adequado do que os modelos RL-AR(1)-N-DI, RL-AR(2)-N-DI, RL-AR(3)-N-DI, RL-AR(1)-T-DI, RL-AR(2)-T-DI, RL-AR(1)-SL-DI, RL-AR(1)-NC-DI, RL-AR(2)-NC-DI e RL-AR(3)-NC-DI. Verificando o fator de Bayes do modelo RL-AR(3)-T-DI com relação aos modelos RL-AR(2)-SL-DI e RL-AR(3)-SL-DI, os resultados 2,146427 e 5,55648, indicam que o modelos mais satisfatórios foram os modelos RL-AR(3)-T-DI, RL-AR(2)-SL-DI, RL-AR(3)-SL-DI, RL-AR(2)-T-DI e RL-AR(1)-N-DI exatamente nesta ordem.

Com o intuito de realizar a análise preditiva, determinamos as estimativas Bayesianas dos parâmetros para os 5 modelos RL-AR($p$)-MEN-DI com os melhores resultados considerando o fator de Bayes, preservando as 12 últimas observações para fins de comparação de previsão e calculamos o erro de previsão quadrático médio (MSPE), também efetuado por Schumacher, Lachos e Dey (2017) e definido como

$$\text{MSPE}_{M_k} = \frac{1}{12}\sum_{t=170}^{181}\left(\log(\text{P}_t) - \widehat{\log(\text{P}_t)}_{M_k}\right)^2,$$

em que $\widehat{\log(\mathrm{P}_t)}_{M_k}$ denota a predição do $t$-ésimo valor do modelo ($M_k$), com $t = 170, \ldots, 181$, dos casos RL-AR($p$)-MEN-DI, definido pela Equação (3.2). Outra medida empregada para mensurar a acurácia do modelo e utilizada como critério para comparação de modelos para previsão é o MAE (Erro Absoluto Médio), calculado a partir do módulo de cada erro para as 12 previsões, como segue:

$$\mathrm{MAE} = \frac{1}{12} \sum_{t=170}^{181} |\log(\mathrm{P}_t) - \widehat{\log(\mathrm{P}_t)}_{M_k}|.$$

**Tabela 7 – MPSE e MAE para os modelos RL-AR($p$)-MEN-DI.**

| **N. ° de valores** | **Modelo** ($M_k$) | **MSPE**$_{M_k}$ | **MAE**$_{M_k}$ |
|---|---|---|---|
| 6 | LR-AR(1)-N-DI | 0,19139 | 0,32673 |
| | **LR-AR(2)-T-DI** | **0,13639** | **0,29796** |
| | LR-AR(3)-T-DI | 0,14914 | 0,31108 |
| | LR-AR(2)-SL-DI | 0,14830 | 0,30504 |
| | LR-AR(3)-SL-DI | 0,15987 | 0,31743 |
| 9 | LR-AR(1)-N-DI | 0,13469 | 0,26576 |
| | **LR-AR(2)-T-DI** | **0,09644** | **0,23852** |
| | LR-AR(3)-T-DI | 0,10486 | 0,24538 |
| | LR-AR(2)-SL-DI | 0,10481 | 0,23959 |
| | LR-AR(3)-SL-DI | 0,11201 | 0,24668 |
| 12 | LR-AR(1)-N-DI | 0,12074 | 0,25326 |
| | **LR-AR(2)-T-DI** | **0,09260** | **0,23798** |
| | LR-AR(3)-T-DI | 0,09893 | 0,24155 |
| | LR-AR(2)-SL-DI | 0,09946 | 0,23820 |
| | LR-AR(3)-SL-DI | 0,10523 | 0,24189 |

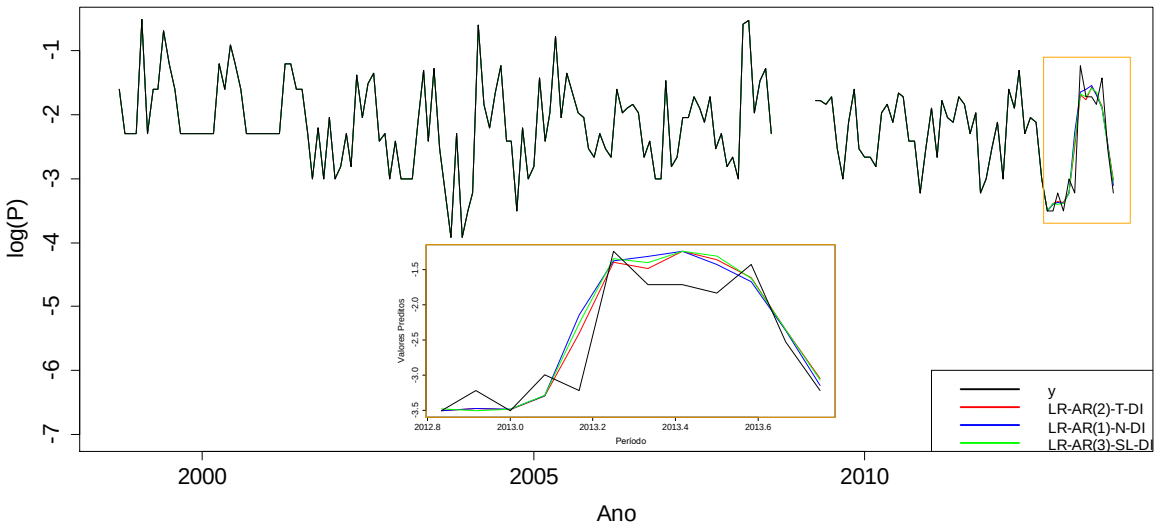
Fonte: Elaborada pelo autor (2022).

Considerando as medidas comparativas MSPE e MAE para os $\{6, 9, 12\}$ passos à frente dos últimos valores para a previsão, apresentados na Tabela 7, os resultados sugerem que o melhor modelo para a previsão é o RL-AR(2)-T-DI, com os menores MPSE e MAE.

Para verificar esse ajuste, apresentamos na Figura 8 a série temporal do logaritmo da concentração de fósforo, com as previsões dos 12 últimos valores para o modelo LR-AR(2)-T-DI, indicado como o melhor modelo de previsão, bem como também para os modelos LR-AR(3)-SL-DI e LR-AR(1)-N-DI, os quais mostraram-se menos satisfatórios para as previsões (Tabela 7), pelo que a observação da figura 8 confirma a consistência e acurácia das previsões.

Conforme visto nos resultados da análise preditiva, percebemos que o modelo RL-AR(2)-T-DI encontra-se nas primeiras posições na Tabela 7 e com melhor ajuste na Figura 8.

**Figura 8 – Série do logaritmo da concentração de fósforo com as previsões sob alguns modelos RL-AR($p$)-MEN-DI.**




Fonte: Elaborada pelo autor (2022).

Neste sentido, como o modelo do caso $t$ de student, com $p = 2$ foi o de maior destaque nas análises, um resumo de medidas das estimativas a posteriori dos parâmetros para o modelo RL-AR(2)-T-DI estão apresentados na Tabela 8.

**Tabela 8 – Estimativas Bayesianas dos parâmetros para o modelo RL-AR(2)-T-DI.**

| **Parâmetro** | **Média** | **Desvio Padrão** | **Percentil** $2,5\%$ | **Percentil** $97,5\%$ |
|---|---|---|---|---|
| $\beta_{1,1}$ | -4,35457 | 0,49799 | -5,36437 | -3,40324 |
| $\beta_{1,2}$ | -2,77692 | 0,74789 | -4,30027 | -1,37201 |
| $\beta_{1,3}$ | -4,25773 | 0,41818 | -5,09446 | -3,44319 |
| $\beta_{1,4}$ | -5,00528 | 0,49410 | -5,98946 | -4,04309 |
| $\beta_{2,1}$ | 0,28681 | 0,08906 | 0,11826 | 0,46723 |
| $\beta_{2,2}$ | 0,14259 | 0,10386 | -0,05289 | 0,35482 |
| $\beta_{2,3}$ | 0,38395 | 0,06905 | 0,24887 | 0,52156 |
| $\beta_{2,4}$ | 0,41735 | 0,09182 | 0,23898 | 0,60076 |
| $\phi_1$ | -0,05351 | 0,08202 | -0,18801 | 0,24155 |
| $\phi_2$ | 0,11753 | 0,07450 | -0,02525 | 0,26671 |
| $\sigma^2$ | 0,15622 | 0,03367 | 0,09752 | 0,23001 |
| $\nu$ | 3,82777 | 1,36467 | 2,05225 | 7,18662 |

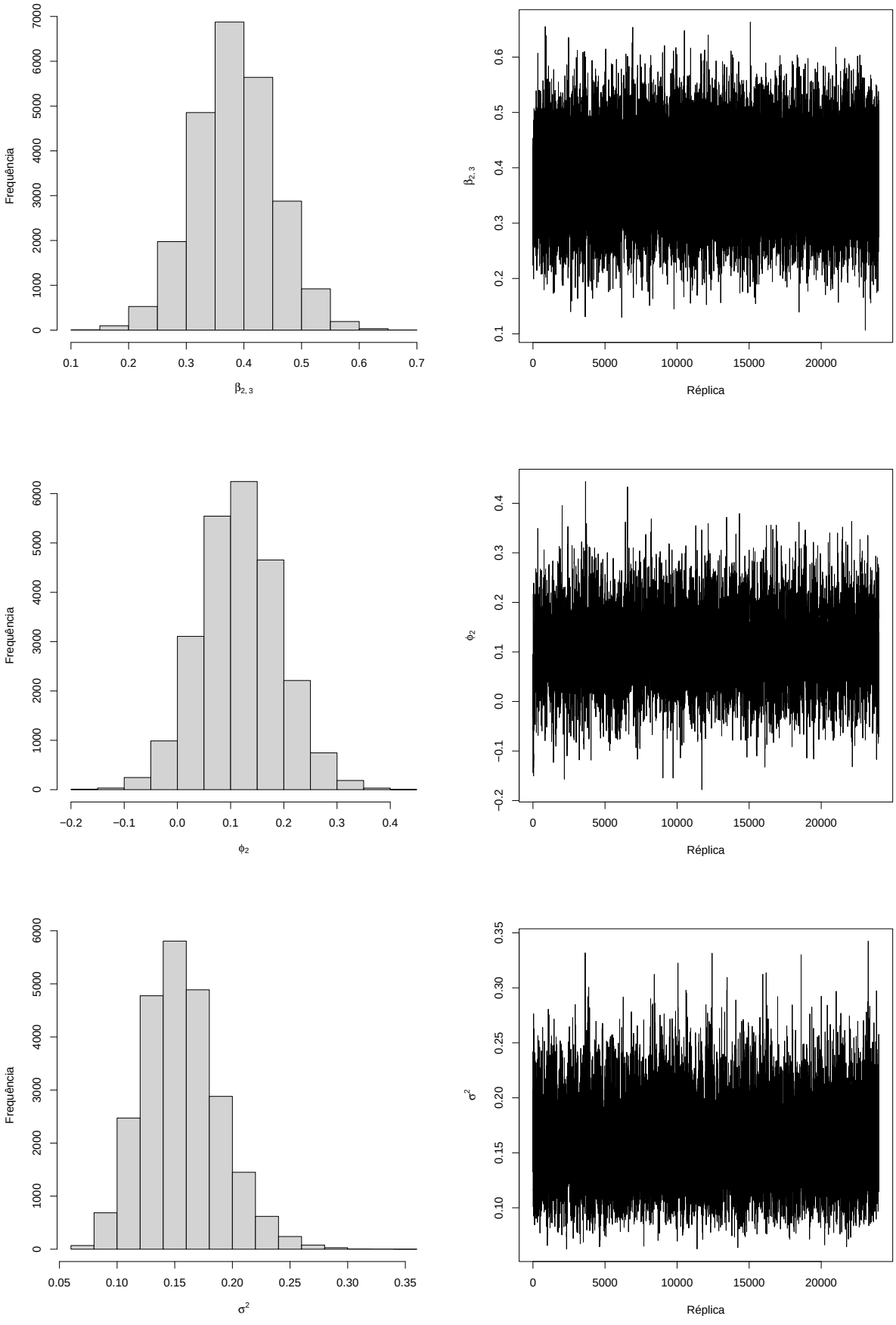
Fonte: Elaborada pelo autor (2022).

A partir da observação da Tabela 8, verificamos, em geral, que a maioria dos parâmetros foram significativos, notamos também a influência das variáveis regressoras no modelo, aquelas que auxiliariam na redução ou no aumento do logaritmo da concentração de fósforo. Os parâmetros $\beta_{1,2}$ e $\beta_{2,2}$ foram os que apresentaram as menores influências, enquanto

que os parâmetros $\beta_{1,4}$ e $\beta_{2,4}$ foram os que apresentaram as maiores influências.

De forma a complementar os resultados dos parâmetros sob o modelo RL-AR(2)-T-DI, apresentamos o histograma da cadeia final, com a ideia de demonstrar a forma aproximada da densidade marginal a posteriori e histórico da cadeia de alguns parâmetros, conforme exposto na Figura 9 . Caso haja interesse na visualização dos demais resultados, estes estão disponíveis com os autores deste trabalho.

**Figura 9 – Histograma e histórico da cadeia final dos parâmetros $\beta_{2,3}, \phi_2$ e $\sigma^2$ sob o modelo RL-AR(2)-T-DI.**




Fonte: Elaborada pelo autor (2022).

# 6 CONSIDERAÇÕES FINAIS

Neste trabalho, propomos um modelo de regressão autoregressivo de ordem $p$ com inovações seguindo uma distribuição na classe misturas da escala normal para dados incompletos, denotado por RL-AR($p$)-MEN-DI, devido a presença de censura (à esquerda, à direita ou intervalar) e/ou *missing data* na variável resposta.

Desenvolvemos um algoritmo MCMC e ajustamos ao modelo proposto, afim de verificar o comportamento da estimação Bayesiana, que foi analisado por meio de três estudos de simulação e pela aplicação a dados reais.

Comparamos os modelos com dados censurados e/ou dados faltantes nas simulações realizadas. Constatamos que o modelo proposto pode fornecer estimativas confiáveis a partir de um conjunto de dados incompletos, com diferentes porcentagens de valores com censura e de valores faltantes, mostrando ainda que as diferenças de inferência entre as abordagens podem ser sutis, sendo, no entanto, perceptível a melhora de ajuste do modelo ao conjunto de dados com valores faltantes.

A eficácia do algoritmo proposto foi avaliado a partir da análise de um conjunto de dados reais, com uma análise comparativa e preditiva dos modelos, indicando que o modelo do caso $t$ de student foi aquele que melhor se ajustou aos dados e o mais adequado na realização das previsões.

Por fim, acreditamos que a partir dos resultados deste trabalho outras pesquisas derivadas podem ser propostas, podendo seus resultados ser estendidos tanto ao modelo AR multivariado ou a modelos ARMA, assim como também a outros tipos de modelos.

# REFERÊNCIAS


ALBERT, J. **Bayesian computation with R**. 1. ed. New York: Springer, 2007.

ANDREWS, D. F.; MALLOWS, C. L. Scale mixtures of normal distributions. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 36, n. 1, p. 99–102, 1974.

BARNDORFF-NIELSEN, O.; SCHOU, G. On the parametrization of autoregressive models by partial autocorrelations. **Journal of multivariate Analysis**, Elsevier, v. 3, n. 4, p. 408–419, 1973.

BOX, G. E.; TIAO, G. C. **Bayesian inference in statistical analysis**. 1. ed. New Jersey: John Wiley & Sons, 1992.

CABRAL, C. R. B.; LACHOS, V. H.; MADRUGA, M. R. Bayesian analysis of skew-normal independent linear mixed models with heterogeneity in the random-effects population. **Journal of Statistical Planning and Inference**, Elsevier, v. 142, n. 1, p. 181–200, 2012.

CASELLA, G.; BERGER, R. L. **Statistical inference**. 2. ed. Connecticut: Cengage Learning, 2010.

CHIB, S. Bayes regression with autoregressive errors: A gibbs sampling approach. **Journal of econometrics**, Elsevier, v. 58, n. 3, p. 275–294, 1993.

CHRISTMAS, J.; EVERSON, R. Robust autoregression: Student-t innovations using variational bayes. **IEEE Transactions on Signal Processing**, IEEE, v. 59, n. 1, p. 48–57, 2010.

CONGDON, P. **Bayesian statistical modelling**. 2. ed. [S.l.]: John Wiley & Sons, 2007.

EHLERS, R. S. Análise de séries temporais. **Laboratório de Estatística e Geoinformação. Universidade Federal do Paraná**, v. 1, p. 1–118, 2007.

FONSECA, T. C.; FERREIRA, M. A.; MIGON, H. S. Objective bayesian analysis for the student-t regression model. **Biometrika**, Oxford University Press, v. 95, n. 2, p. 325–333, 2008.

GAMERMAN, D.; LOPES, H. F. **Markov chain Monte Carlo: stochastic simulation for Bayesian inference**. Boca Raton: CRC press, 2006.

GARAY, A. M.; BOLFARINE, H.; LACHOS, V. H.; CABRAL, C. R. Bayesian analysis of censored linear regression models with scale mixtures of normal distributions. **Journal of Applied Statistics**, Taylor & Francis, v. 42, n. 12, p. 2694–2714, 2015.

GARAY, A. M.; LACHOS, V. H.; BOLFARINE, H.; CABRAL, C. R. Linear censored regression models with scale mixtures of normal distributions. **Statistical Papers**, Springer, v. 58, n. 1, p. 247–278, 2017.

GARAY, A. M.; MEDINA, F. L.; CABRAL, C. R.; LIN, T.-I. Bayesian analysis of the p-order integer-valued ar process with zero-inflated poisson innovations. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 90, n. 11, p. 1943–1964, 2020.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. Bayesian data analysis chapman & hall. **CRC Texts in Statistical Science**, 2004.

GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, n. 6, p. 721–741, 1984.

GEWEKE, J. Bayesian treatment of the independent student-t linear model. **Journal of applied econometrics**, Wiley Online Library, v. 8, n. S1, p. S19–S40, 1993.

GHASAMI, S.; KHODADADI, Z.; MALEKI, M. Autoregressive processes with generalized hyperbolic innovations. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, v. 49, n. 12, p. 3080–3092, 2020.

GHASAMI, S.; MALEKI, M.; KHODADADI, Z. Leptokurtic and platykurtic class of robust symmetrical and asymmetrical time series models. **Journal of Computational and Applied Mathematics**, Elsevier, v. 376, p. 112806, 2020.

GILKS, W.; RICHARDSON, S.; SPIEGELHALTER, D. **Markov chain Monte Carlo in practice**. 1. ed. New York: Chapman and Hall/CRC, 1996.

HAGHBIN, H.; NEMATOLLAHI, A. Likelihood-based inference in autoregressive models with scaled t-distributed innovations by means of em-based algorithms. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, v. 42, n. 10, p. 2239–2252, 2013.

HAMILTON, J. **Time Series Analysis, New Jersey, Princeton Uni**. [S.l.]: Press, 1994.

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

JEFFREYS, H. **Theory of probability**. 3. ed. Oxford: UK: Claredon, 1998.

KOUL, H. L.; DESHPANDE, J. V. Analysis of censored data: Proceedings of the workshop on analysis of censored data, december 28, 1994-january 1, 1995. In: IMS. University of Pune, Pune, India, 1995.

LIBRA, R. D.; WOLTER, C. F.; LANGEL, R. J. Nitrogen and phosphorus budgets for iowa and iowa watersheds. Iowa Department of Natural Resources, Geological Survey, 2004.

LIN, T. I.; LEE, J. C. Bayesian analysis of hierarchical linear mixed modeling using the multivariate t distribution. **Journal of Statistical Planning and Inference**, Elsevier, v. 137, n. 2, p. 484–495, 2007.

LITTLE, R. J.; RUBIN, D. B. **Statistical analysis with missing data**. Hoboken: John Wiley & Sons, 2019. v. 793.

LIU, J.; KUMAR, S.; PALOMAR, D. P. Parameter estimation of heavy-tailed ar model with missing data via stochastic em. **IEEE Transactions on Signal Processing**, IEEE, v. 67, n. 8, p. 2159–2172, 2019.

LIU, J. S. The collapsed gibbs sampler in bayesian computations with applications to a gene regulation problem. **Journal of the American Statistical Association**, Taylor & Francis, v. 89, n. 427, p. 958–966, 1994.

MALEKI, M.; ARELLANO-VALLE, R. B.; DEY, D. K.; MAHMOUDI, M. R.; JALALI, S. M. J. A bayesian approach to robust skewed autoregressive processes. **Calcutta Statistical Association Bulletin**, SAGE Publications Sage India: New Delhi, India, v. 69, n. 2, p. 165–182, 2017.

MALEKI, M.; NEMATOLLAHI, A. Autoregressive models with mixture of scale mixtures of gaussian innovations. **Iranian Journal of Science and Technology, Transactions A: Science**, Springer, v. 41, n. 4, p. 1099–1107, 2017.

MASSUIA, M. B.; GARAY, A. M.; CABRAL, C. R.; LACHOS, V. Bayesian analysis of censored linear regression models with scale mixtures of skew-normal distributions. **Statistics and its Interface**, International Press of Boston, v. 10, n. 3, p. 425–439, 2017.

MCKNIGHT, P. E.; MCKNIGHT, K. M.; SIDANI, S.; FIGUEREDO, A. J. **Missing data: A gentle introduction**. New York: Guilford Press, 2007.

MCKNIGHT, S. D.; MCKEAN, J. W.; HUITEMA, B. E. A double bootstrap method to analyze linear models with autoregressive error terms. **Psychological methods**, American Psychological Association, v. 5, n. 1, p. 87, 2000.

METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. **The journal of chemical physics**, American Institute of Physics, v. 21, n. 6, p. 1087–1092, 1953.

NDUKA, U. C. Em-based algorithms for autoregressive models with t-distributed innovations. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, v. 47, n. 1, p. 206–228, 2018.

NTZOUFRAS, I. **Bayesian modeling using WinBUGS**. Hoboken: John Wiley & Sons, 2009.

OHTANI, K. On estimating and testing in a linear regression model with autocorrelated errors. **Journal of Econometrics**, Elsevier, v. 44, n. 3, p. 333–346, 1990.

PRADO, R.; WEST, M. **Time series: modeling, computation, and inference**. [S.l.]: Chapman and Hall/CRC, 2010.

RIZZO, M. L. **Statistical computing with R**. London: Chapman and Hall/CRC, 2007.

ROSADI, D.; FILZMOSER, P. Robust second-order least-squares estimation for regression models with autoregressive errors. **Statistical Papers**, Springer, v. 60, n. 1, p. 105–122, 2019.

SCHUMACHER, F. L.; LACHOS, V. H.; DEY, D. K. Censored regression models with autoregressive errors: A likelihood-based perspective. **Canadian Journal of Statistics**, Wiley Online Library, v. 45, n. 4, p. 375–392, 2017.

SHIN, D. W.; SARKAR, S. Parameter estimation in regression models with autocorrelated errors using irregular data. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 23, n. 12, p. 3567–3580, 1994.

SHUMWAY, R. H.; STOFFER, D. S. **Time series analysis and its applications**. [S.l.]: Springer, 2000. v. 3.

TARAMI, B.; POURAHMADI, M. Multi-variate t autoregressions: Innovations, prediction variances and exact likelihood equations. **Journal of Time Series Analysis**, Wiley Online Library, v. 24, n. 6, p. 739–754, 2003.

THOMSON, T.; HOSSAIN, S.; GHAHRAMANI, M. Application of shrinkage estimation in linear regression models with autoregressive errors. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 85, n. 16, p. 3335–3351, 2015.

TIKU, M. L.; WONG, W.-K.; VAUGHAN, D. C.; BIAN, G. Time series models in non-normal situations: Symmetric innovations. **Journal of Time Series Analysis**, Wiley Online Library, v. 21, n. 5, p. 571–596, 2000.

VALERIANO, K. A.; SCHUMACHER, F. L.; GALARZA, C. E.; MATOS, L. A. Censored autoregressive regression models with student-*t* innovations. **arXiv preprint arXiv:2110.00224**, 2021.

WANG, C.; CHAN, K.-S. Quasi-likelihood estimation of a censored autoregressive model with exogenous variables. **Journal of the American Statistical Association**, Taylor & Francis, v. 113, n. 523, p. 1135–1145, 2018.

WOOD, A. M.; WHITE, I. R.; THOMPSON, S. G. Are missing outcome data adequately handled? a review of published randomized controlled trials in major medical journals. **Clinical trials**, Sage Publications Sage CA: Thousand Oaks, CA, v. 1, n. 4, p. 368–376, 2004.

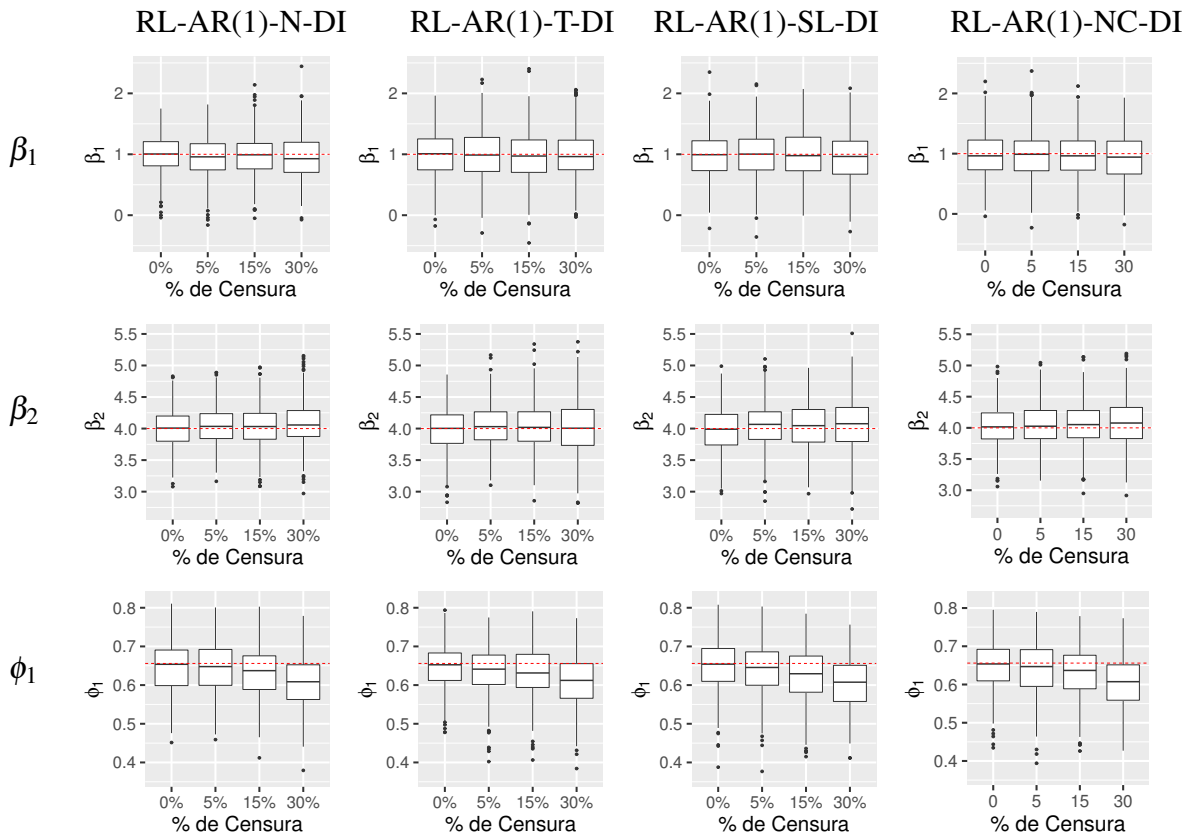
ZEGER, S. L.; BROOKMEYER, R. Regression analysis with censored autocorrelated data. **Journal of the American Statistical Association**, Taylor & Francis, v. 81, n. 395, p. 722–729, 1986.

ZHOU, R.; LIU, J.; KUMAR, S.; PALOMAR, D. P. Parameter estimation for student'st var model with missing data. In: IEEE. **ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. [S.l.], 2021. p. 5145–5149.

# APÊNDICE A - RESULTADOS COMPLEMENTARES DO CENÁRIO 1 DOS ESTUDOS DE SIMULAÇÃO

Neste apêndice apresentamos os resultados do Cenário 1 dos Estudos de Simulação para $T = 150$ e diferentes níveis de censura NCens $\in \{0\%, 5\%, 15\%, 30\%\}$.

**Figura A.1 – Estimativas bayesianas médias de $\beta_1$, $\beta_2$ e $\phi_1$ para os modelos RL-AR(1)-MEN-DI, com $T = 150$ e diferente níveis de censura.**



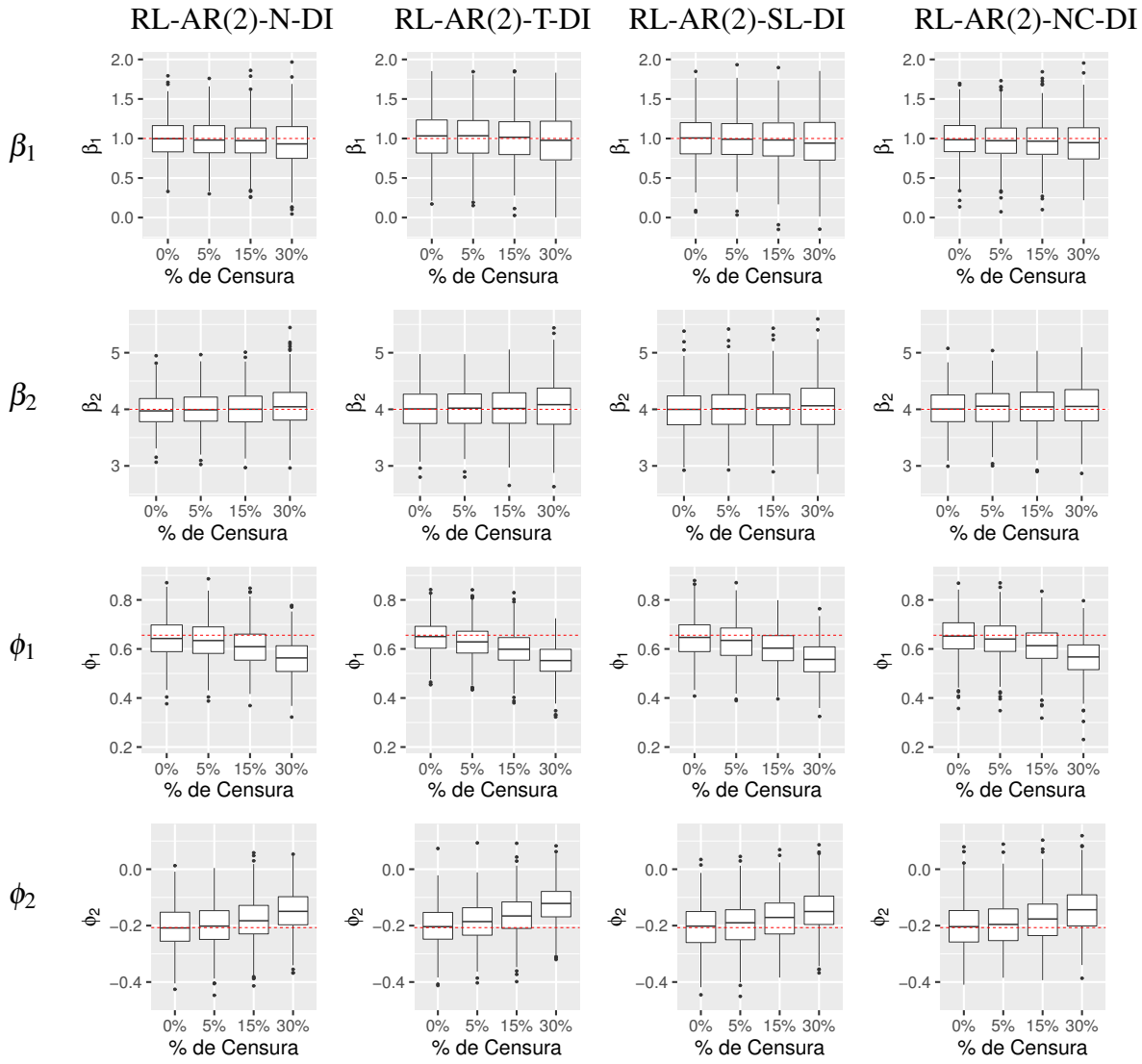

Fonte: Elaborada pelo autor (2022).

**Tabela A.1 – Resumo das estimativas bayesianas baseadas nas 500 réplicas simuladas dos modelos RL-AR($p$)-MEN-DI, para $p = 1, T = 150$ e diferentes níveis de censura.**

| | | RL-AR(1)-N-DI | | | RL-AR(1)-T-DI | | | RL-AR(1)-SL-DI | | | RL-AR(1)-NC-DI | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **NCens** | **PR** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** |
| 0% | $\beta_1$ | 1,0055 | 0,3646 | 95,4% | 0,9968 | 0,4046 | 96,6% | 0,9758 | 0,4213 | 95,8% | 0,9835 | 0,3914 | 96,6% |
| | $\beta_2$ | 4,0053 | 0,2965 | 93,2% | 3,9956 | 0,3507 | 94,8% | 3,9816 | 0,3361 | 95,4% | 4,0213 | 0,3216 | 95,4% |
| | $\phi_1$ | 0,6467 | 0,0653 | 95,6% | 0,6471 | 0,0554 | 97,0% | 0,6492 | 0,0646 | 94,0% | 0,6491 | 0,0652 | 95,8% |
| | $\sigma^2$ | 1,5324 | 0,1819 | 95,2% | 1,4599 | 0,2850 | 93,8% | 1,5081 | 0,2965 | 99,0% | 1,3636 | 0,4139 | 98,8% |
| 5% | $\beta_1$ | 0,9514 | 0,3657 | 96,2% | 0,9970 | 0,4110 | 94,2% | 0,9990 | 0,4234 | 93,6% | 0,9779 | 0,3956 | 94,8% |
| | $\beta_2$ | 4,0439 | 0,3013 | 95,0% | 4,0388 | 0,3528 | 95,2% | 4,0418 | 0,3545 | 95,0% | 4,0435 | 0,3307 | 93,6% |
| | $\phi_1$ | 0,6435 | 0,0661 | 95,0% | 0,6388 | 0,0598 | 95,6% | 0,6408 | 0,0675 | 95,4% | 0,6412 | 0,0667 | 96,2% |
| | $\sigma^2$ | 1,5462 | 0,1931 | 96,0% | 1,4903 | 0,3044 | 96,0% | 1,5409 | 0,3132 | 99,6% | 1,4021 | 0,4091 | 99,4% |
| 15% | $\beta_1$ | 0,9826 | 0,3692 | 95,0% | 0,9693 | 0,4477 | 95,4% | 0,9972 | 0,4155 | 94,6% | 0,9698 | 0,3974 | 96,6% |
| | $\beta_2$ | 4,0363 | 0,3261 | 94,8% | 4,0236 | 0,3801 | 95,6% | 4,0351 | 0,3786 | 97,0% | 4,0600 | 0,3564 | 95,4% |
| | $\phi_1$ | 0,6316 | 0,0689 | 95,8% | 0,6340 | 0,0636 | 94,0% | 0,6245 | 0,0710 | 94,4% | 0,6310 | 0,0695 | 96,8% |
| | $\sigma^2$ | 1,6175 | 0,2257 | 93,6% | 1,5485 | 0,3378 | 96,6% | 1,5669 | 0,3309 | 99,2% | 1,4880 | 0,4086 | 99,4% |
| 30% | $\beta_1$ | 0,9469 | 0,4033 | 96,4% | 0,9873 | 0,4837 | 97,0% | 0,9565 | 0,4464 | 95,6% | 0,9397 | 0,4250 | 96,2% |
| | $\beta_2$ | 4,0773 | 0,4015 | 97,0% | 4,0178 | 0,4423 | 95,8% | 4,0746 | 0,4490 | 97,6% | 4,0866 | 0,4303 | 97,6% |
| | $\phi_1$ | 0,6095 | 0,0747 | 96,0% | 0,6084 | 0,0703 | 92,4% | 0,6030 | 0,0766 | 94,6% | 0,6032 | 0,0757 | 93,8% |
| | $\sigma^2$ | 1,7619 | 0,2980 | 86,4% | 1,6544 | 0,4068 | 96,2% | 1,7335 | 0,4046 | 99,2% | 1,6214 | 0,4473 | 99,8% |

Fonte: Elaborada pelo autor (2022).
Simulação 01. Nível de Censura (NCens), Parâmetro (PR), Média das médias das 500 réplicas (M Med), Média dos desvios padrão das 500 réplicas (M DP), Percentual de cobertura do intervalo de credibilidade de 95% (P Cob).

**Figura A.2 – Estimativas bayesianas médias de $\beta_1$, $\beta_2$, $\phi_1$ e $\phi_2$ para os modelos RL-AR(2)-MEN-DI, com $T = 150$ e diferente níveis de censura.**




Fonte: Elaborada pelo autor (2022).

**Tabela A.2 – Resumo das estimativas bayesianas baseadas nas 500 réplicas simuladas dos modelos RL-AR($p$)-MEN-DI, para $p = 2, T = 150$ e diferentes níveis de censura.**

| | | RL-AR(2)-N-DI | | | RL-AR(2)-T-DI | | | RL-AR(2)-SL-DI | | | RL-AR(2)-NC-DI | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **NCens** | **PR** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** | **M Med** | **M DP** | **P Cob** |
| 0% | $\beta_1$ | 1,0015 | 0,2498 | 94,8% | 1,0172 | 0,2946 | 94,2% | 1,0100 | 0,2864 | 95,6% | 0,9951 | 0,2713 | 95,8% |
| | $\beta_2$ | 3,9814 | 0,3066 | 95,2% | 4,0044 | 0,3631 | 94,2% | 3,9746 | 0,3509 | 93,2% | 4,0078 | 0,3286 | 95,8% |
| | $\phi_1$ | 0,6413 | 0,0820 | 93,6% | 0,6465 | 0,0700 | 95,2% | 0,6454 | 0,0816 | 96,2% | 0,6504 | 0,0819 | 94,6% |
| | $\phi_2$ | -0,2028 | 0,0822 | 95,8% | -0,2011 | 0,0698 | 94,6% | -0,2028 | 0,0817 | 95,8% | -0,2013 | 0,0821 | 94,2% |
| | $\sigma^2$ | 1,5172 | 0,1815 | 94,2% | 1,4695 | 0,2902 | 92,8% | 1,5029 | 0,2967 | 99,2% | 1,3440 | 0,4140 | 99,0% |
| 5% | $\beta_1$ | 0,9873 | 0,2563 | 95,4% | 1,0184 | 0,3028 | 95,2% | 1,0011 | 0,2952 | 96,2% | 0,9792 | 0,2808 | 95,6% |
| | $\beta_2$ | 3,9986 | 0,3173 | 96,2% | 4,0121 | 0,3766 | 95,2% | 3,9910 | 0,3642 | 94,8% | 4,0296 | 0,3408 | 96,0% |
| | $\phi_1$ | 0,6332 | 0,0837 | 93,6% | 0,6270 | 0,0767 | 93,8% | 0,6316 | 0,0854 | 95,6% | 0,6399 | 0,0842 | 95,4% |
| | $\phi_2$ | -0,1969 | 0,0839 | 96,0% | -0,1858 | 0,0735 | 96,0% | -0,1931 | 0,0848 | 96,8% | -0,1942 | 0,0841 | 95,6% |
| | $\sigma^2$ | 1,5541 | 0,1957 | 94,4% | 1,4948 | 0,3083 | 95,2% | 1,5240 | 0,3138 | 99,4% | 1,4004 | 0,4061 | 99,6% |
| 15% | $\beta_1$ | 0,9782 | 0,2741 | 96,2% | 1,0019 | 0,3197 | 95,2% | 0,9886 | 0,3213 | 97,0% | 0,9662 | 0,2974 | 97,0% |
| | $\beta_2$ | 4,0079 | 0,3456 | 96,6% | 4,0271 | 0,4052 | 95,8% | 4,0058 | 0,3962 | 94,8% | 4,0432 | 0,3733 | 96,8% |
| | $\phi_1$ | 0,6082 | 0,0887 | 91,4% | 0,5983 | 0,0827 | 92,2% | 0,6018 | 0,0908 | 93,0% | 0,6129 | 0,0895 | 93,0% |
| | $\phi_2$ | -0,1788 | 0,0889 | 95,2% | -0,1642 | 0,0776 | 93,6% | -0,1733 | 0,0897 | 95,8% | -0,1763 | 0,0892 | 96,0% |
| | $\sigma^2$ | 1,6103 | 0,2268 | 94,4% | 1,5730 | 0,3423 | 96,4% | 1,5957 | 0,3415 | 99,0% | 1,5080 | 1,3215 | 99,0% |
| 30% | $\beta_1$ | 0,9401 | 0,3230 | 95,4% | 0,9752 | 0,3734 | 96,1% | 0,9511 | 0,3692 | 95,4% | 0,9356 | 0,3491 | 96,8% |
| | $\beta_2$ | 4,0576 | 0,4113 | 96,0% | 4,0647 | 0,4761 | 96,3% | 4,0532 | 0,4659 | 95,2% | 4,0716 | 0,4420 | 96,8% |
| | $\phi_1$ | 0,5615 | 0,0973 | 90,8% | 0,5516 | 0,0912 | 89,0% | 0,5574 | 0,0984 | 90,2% | 0,5669 | 0,0975 | 92,2% |
| | $\phi_2$ | -0,1474 | 0,0977 | 95,2% | -0,1245 | 0,0844 | 92,0% | -0,1466 | 0,0972 | 96,0% | -0,1455 | 0,0971 | 96,0% |
| | $\sigma^2$ | 1,7370 | 0,2964 | 89,6% | 1,7504 | 0,4151 | 95,1% | 1,7363 | 0,4080 | 98,0% | 1,6423 | 0,5752 | 98,0% |

Fonte: Elaborada pelo autor (2022).
Simulação 01. Nível de Censura (NCens), Parâmetro (PR), Média das médias das 500 réplicas (M Med), Média dos desvios padrão das 500 réplicas (M DP), Percentual de cobertura do intervalo de credibilidade de 95% (P Cob).