

UNIVERSIDADE FEDERAL DE SANTA MARIA  
CENTRO DE CIÊNCIAS NATURAIS E EXATAS  
CURSO DE ESPECIALIZAÇÃO EM ESTATÍSTICA E MODELAGEM  
QUANTITATIVA

Cristiane Melchior

**ANÁLISE DA EVOLUÇÃO DAS TAXAS DE MORTALIDADE DE  
CRIANÇAS COM ATÉ UM ANO DE IDADE RESIDENTES NO  
ESTADO DO RIO GRANDE DO SUL**

Santa Maria, RS  
2021

**Cristiane Melchior**

**ANÁLISE DA EVOLUÇÃO DAS TAXAS DE MORTALIDADE DE CRIANÇAS COM  
ATÉ UM ANO DE IDADE RESIDENTES NO ESTADO DO RIO GRANDE DO SUL**

Monografia apresentada ao curso de Especialização em Estatística e Modelagem Quantitativa, da Universidade Federal de Santa Maria (UFSM, RS) como requisito parcial para obtenção do título de **Especialista em Estatística e Modelagem Quantitativa.**

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Roselaine Ruviaro Zanini

Santa Maria, RS  
2021

**Cristiane Melchior**

**ANÁLISE DA EVOLUÇÃO DAS TAXAS DE MORTALIDADE DE CRIANÇAS COM  
ATÉ UM ANO DE IDADE RESIDENTES NO ESTADO DO RIO GRANDE DO SUL**

Monografia apresentada ao curso de Especialização em Estatística e Modelagem Quantitativa, da Universidade Federal de Santa Maria (UFSM, RS) como requisito parcial para obtenção do título de **Especialista em Estatística e Modelagem Quantitativa.**

**Aprovado em 11 de janeiro de 2021**



**Roselaine Ruviano Zanini, Dra. (UFSM)**  
(Presidente/Orientadora)



**Luciane Flores Jacobi, Dra. (UFSM)**



**Adriano Mendonça Souza, Dr. (UFSM)**

**Santa Maria, RS  
2021**

## **AGRADECIMENTOS**

*Primeiramente, agradeço à Deus por me conceder saúde, conhecimento e perseverança para superar meus limites.*

*Obrigada aos meus pais Elemar e Edite, meus irmãos Cristian e Douglas e meu namorado Dinei, pelo apoio, paciência e amor nos momentos em que estes se fizeram necessários.*

*Agradeço à minha querida orientadora Professora Dra. Roselaine Ruviano Zanini por todo o auxílio, compreensão e conhecimento. Sou muito grata pela confiança e grande oportunidade de trabalhar ao seu lado, em mais uma etapa da minha vida.*

*Obrigada à banca pelas contribuições para a lapidação e encerramento deste trabalho.*

*Gratidão à Universidade Federal de Santa Maria pela concessão do título de Especialista.*

## RESUMO

### **ANÁLISE DA EVOLUÇÃO DAS TAXAS DE MORTALIDADE DE CRIANÇAS COM ATÉ UM ANO DE IDADE RESIDENTES NO ESTADO DO RIO GRANDE DO SUL**

AUTORA: Cristiane Melchior  
ORIENTADORA: Roselaine Ruviaro Zanini

As taxas de mortalidade infantil representam indicadores de qualidade de vida, além de medir o desenvolvimento humano, a desigualdade social e as condições de vida e saúde de determinada população. Nesse sentido, este estudo tem por propósito destacar as principais características e comparar a evolução das taxas de mortalidade infantil no estado do Rio Grande do Sul (RS) por meio das metodologias ARIMA e  $\beta$ ARMA. As taxas de mortalidade infantil foram obtidas do Departamento de Informática do SUS (DATASUS), por meio dos óbitos registrados no Sistema de Informação sobre Mortalidade (SIM), estratificando-os por estado. De posse dessas informações foi realizada uma análise descritiva e demonstradas as características de vulnerabilidade dos bebês, das mães e do parto. Além disso, foi modelada a série temporal utilizando as metodologias Box-Jenkins ou ARIMA e  $\beta$ ARMA, considerando observações mensais no período de 2000 a 2017. Cacique Doble foi o município de residência com a maior taxa de mortalidade infantil para o estado do RS neste período. Utilizando a análise dos resíduos e os critérios penalizadores AIC e BIC para avaliar a qualidade dos modelos observou-se um melhor desempenho no ajuste do modelo  $\beta$ ARMA, sendo que o melhor modelo para a série foi  $\beta$ ARMA(4,6) no horizonte de tempo de 6 meses. Porém, as medidas de acurácia MAE, MAPE e RMSE foram menores no modelo SARIMA. As metodologias propostas podem orientar no planejamento de políticas preventivas e educativas voltadas ao risco de um nascido vivo morrer durante o seu primeiro ano de vida.

Palavras-chave: Taxa de mortalidade infantil. Previsão. Séries temporais. ARIMA.  $\beta$ ARMA.

## ABSTRACT

### ANALYSIS OF THE EVOLUTION OF MORTALITY RATES OF CHILDREN UP TO ONE YEAR OLD RESIDING IN THE STATE OF RIO GRANDE DO SUL

AUTHOR: Cristiane Melchior  
ADVISOR: Roselaine Ruviaro Zanini

Infant mortality rates represent indicators of quality of life and measure human development, social inequality, and the living and health conditions of a given population. This study aims to highlight their main characteristics and compare the evolution of infant mortality rates in the state of Rio Grande do Sul (RS) through the ARIMA and  $\beta$ ARMA methodologies. The infant mortality rates were obtained from the Unified Health System (SUS) Department of Informatics (DATASUS), through the deaths registered in the Mortality Information System (SIM), stratifying them by state. A descriptive analysis was carried out and the vulnerability characteristics of babies, mothers, and childbirth were demonstrated. In addition, the time series was modeled using the Box-Jenkins or ARIMA and  $\beta$ ARMA methodologies, considering monthly observations in the period of 2000 to 2017. Cacique Doble was the city of residence with the highest infant mortality rate for the state of RS registered in this period. Using the residual analysis and the AIC and BIC penalizing criteria to evaluate the quality of the models, a better performance was observed in the adjustment of the  $\beta$ ARMA model, while the best model for the series was  $\beta$ ARMA(4,6) in the 6-month time horizon. However, the accuracy measures MAE, MAPE and RMSE were lower in the SARIMA model. The proposed methodologies can guide the planning of preventive and educational policies aimed at the risk of a born alive dying during its first year of life.

Keywords: Child mortality rate. Prediction. Time series. ARIMA.  $\beta$ ARMA.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>6</b>
1.1	PROBLEMA.....	7
1.2	OBJETIVOS .....	7
1.2.1	<b>Objetivo Geral</b> .....	<b>7</b>
1.2.2	<b>Objetivos Específicos</b> .....	<b>7</b>
1.3	JUSTIFICATIVA .....	8
1.4	ESTRUTURA DO TRABALHO .....	9
<b>2</b>	<b>REVISÃO DA LITERATURA</b> .....	<b>10</b>
2.1	MORTALIDADE INFANTIL.....	10
<b>3</b>	<b>MATERIAIS E MÉTODOS</b> .....	<b>14</b>
3.1	METODOLOGIA DE BOX-JENKINS .....	15
3.2	METODOLOGIA $\beta$ ARMA .....	19
<b>4</b>	<b>RESULTADOS E DISCUSSÕES</b> .....	<b>24</b>
<b>5</b>	<b>CONCLUSÃO</b> .....	<b>33</b>
	<b>REFERÊNCIAS</b> .....	<b>35</b>

## 1 INTRODUÇÃO

A mortalidade infantil é caracterizada pela quantidade de óbitos ocorridos desde o nascimento até 365 dias após o nascimento em uma certa população (MOHAMMADI *et al.*, 2018; ZYLBERSZTEJN *et al.*, 2018). Estimar o coeficiente de mortalidade infantil tem sido incentivado pelo crescente interesse na mensuração das taxas de óbitos infantis, não apenas como indicador de qualidade da saúde, mas sobretudo como um modo de nivelar o desenvolvimento humano de determinada população (ALMEIDA e SZWARCOWALD, 2014). Nesse contexto, é válido ressaltar que a mortalidade infantil serve como indicador das condições de vida e saúde de um determinado grupo (NASCIMENTO *et al.*, 2014).

As causas atribuídas aos óbitos infantis estão relacionadas diretamente a problemas sociais, condições precárias de domicílio, transporte, falhas ou atrasos na obtenção de atendimento na unidade de saúde, além da contribuição de fatores não biológicos (MAKAROVA *et al.*, 2016).

Entretanto, o modelo de desenvolvimento social e econômico que vigorou no Brasil durante décadas foi restritivo e centralizador de renda, recursos e serviços, favorecendo determinadas regiões e grupos sociais. Por consequência, têm-se uma sociedade marcada pelos princípios da desigualdade social (ALMEIDA e SZWARCOWALD, 2014). Esse modelo de gestão também impulsionou o aumento dos óbitos infantis, reflexo das limitações existentes nas políticas de saúde e condições de vida da população (CARVALHO *et al.*, 2018).

Considerando o elevado índice de mortalidade infantil e as consequências sociais e econômicas destes óbitos (CARVALHO *et al.*, 2018), se torna pertinente realizar uma abordagem com a utilização de metodologias preditivas em séries temporais.

O uso de séries temporais para a observação de eventos espaçados de forma regular no tempo foi desenvolvido por Box e Jenkins (1976), a partir de estudos com ênfase em aplicações no campo da estatística, economia, engenharia e ciências sociais. Os modelos de previsão avaliam dados por meio da análise de séries temporais e atendem a alguns pressupostos por meio da investigação do fator gerador da série temporal, realização de previsões futuras para a série em questão, descrição

do comportamento da série, além de indicar a periodicidade dos dados (MORETTIN e TOLOI, 2006).

Portanto, o objetivo deste estudo é comparar a evolução das taxas de mortalidade infantil no estado do Rio Grande do Sul (RS), por meio das metodologias ARIMA e  $\beta$ ARMA, além de destacar suas principais características.

## 1.1 PROBLEMA

Por se tratar de um fator sensível às privações e alterações sociodemográficas, a mortalidade infantil expressa o desenvolvimento de uma determinada região, suas condições de saúde e disparidades. Prever as futuras taxas de mortalidade infantil releva a seriedade do problema que estas representam. Portanto, qual a metodologia apropriada para estabelecer o melhor modelo de previsão das taxas de mortalidade infantil?

Esta questão foi respondida por meio da análise da série histórica, que apontou indicadores de mortalidade infantil com a finalidade de incentivar o desenvolvimento de políticas e ações de saúde com ênfase no pré-natal, no parto e na proteção da saúde infantil.

## 1.2 OBJETIVOS

Nesta seção são apresentados os objetivos geral e específicos do trabalho.

### 1.2.1 Objetivo Geral

- Comparar a evolução das taxas de mortalidade infantil no estado do RS, por meio das metodologias ARIMA e  $\beta$ ARMA, além de destacar suas principais características

### 1.2.2 Objetivos Específicos

- Descrever variáveis a serem analisadas estratificando-as por sexo, idade, raça/cor, município de residência, entre outras covariáveis disponíveis.

- Utilizar a metodologia ARIMA e  $\beta$ ARMA para modelar as séries de dados referentes às taxas de mortalidade infantil
- Comparar os modelos de previsão ajustados e identificar o melhor modelo preditivo para o estado do RS.

### 1.3 JUSTIFICATIVA

A análise estatística com modelos de previsão pode ser utilizada para auxiliar na mensuração de ocorrências de óbitos infantis (NAVARRO *et al.*, 2013), indicar características de vulnerabilidade, local de residência, a qualidade dos serviços de saúde e saneamento básico, o nível de escolaridade, idade materna, e de renda (MOHAMMADI *et al.*, 2018), dentre outras causas subjacentes da mortalidade infantil (MAKAROVA *et al.*, 2016; ÖZDEMIR *et al.*, 2015).

A estatística é uma ferramenta que fornece métodos e auxilia na interpretação e na análise de dados utilizados na tomada de decisão. É apresentada como um instrumento de coleta, classificação, disposição e interpretação numérica (ADEDINI *et al.*, 2015). Em relação à análise dos índices de mortalidade infantil, Marshall Junior, et al. (2012) ressaltam que a estatística oferece ferramentas para determinar as variáveis de interesse, elaborar tabelas, calcular níveis de confiabilidade, realizar previsões acertadas, análises de regressão, de variância e multivariadas, dentre outras.

No tocante aos modelos de previsão, estes tornam-se relevantes visto que oferecem uma previsão futura de ocorrências do fator estudado (BOX e JENKINS, 1976; MORETTIN e TOLOI, 2006). Dentre os métodos ARIMA e  $\beta$ ARMA, que foram utilizados para a análise da série temporal dos óbitos infantis do estado do RS, foi identificado o modelo previsor mais eficiente.

O presente estudo é fundamental para o entendimento da relação entre as variáveis estudadas e as políticas de saúde, uma vez que as pesquisas existentes não abordam as taxas de mortalidade infantil no estado do RS sob o enfoque das metodologias preditivas ARIMA e  $\beta$ ARMA.

Nesse sentido, as entidades de saúde no estado do RS poderão utilizar-se dos resultados dessa pesquisa para minimizar as características adversas no momento do nascimento e os fatores socioeconômicos responsáveis por essas mortes. Assim, podem ser aplicadas políticas preventivas que abordem as causas de mortalidade

infantil, melhorando a saúde materna antes e durante a gravidez (ZYLBERSZTEJN *et al.*, 2018).

#### 1.4 ESTRUTURA DO TRABALHO

Este estudo está estruturado em cinco capítulos, distribuídos da seguinte forma:

O **primeiro capítulo** descreve a introdução ao tema de pesquisa, o problema, os objetivos da pesquisa, a justificativa e a estrutura do trabalho.

No **segundo capítulo** está a revisão de literatura, que abrange o tópico: mortalidade infantil.

Enquanto que o **terceiro capítulo** descreve os materiais e métodos utilizados, abordando a metodologia de Box-Jenkins que apresenta os modelos estacionários, modelo autorregressivos (AR), modelo de médias móveis (MA), modelo autorregressivo de médias móveis (ARMA), modelos não-estacionários e modelos sazonais. Ao final do capítulo, ainda se apresentam as etapas metodológicas da metodologia Box-Jenkins.

Além disso, neste capítulo também está descrita a metodologia  $\beta$ ARMA que se baseia na distribuição Beta apresentando algumas particularidades que foram evidenciadas nessa terceira seção.

No **quarto capítulo** encontram-se os resultados e discussões.

Enquanto que, no **quinto capítulo** apresentou-se a conclusão deste estudo.

## 2 REVISÃO DA LITERATURA

Neste capítulo é apresentada uma revisão de literatura sobre a mortalidade infantil.

### 2.1 MORTALIDADE INFANTIL

A mortalidade infantil é caracterizada pela ocorrência do óbito entre as crianças menores de um ano de idade, sendo considerada a base para o cálculo da taxa de mortalidade infantil. Esta taxa é gerada a partir da taxa de mortalidade infantil observada durante um ano, considerando o número de nascidos vivos no mesmo período de tempo (MOHAMMADI *et al.*, 2018; ZYLBERSZTEJN *et al.*, 2018). Esse cálculo se dá por meio da equação (1) em determinado espaço geográfico e no ano considerado:

*Coefficiente de mortalidade infantil:*

$$\frac{\text{N}^{\circ} \text{ de óbitos de residentes com menos de um ano de idade}}{\text{N}^{\circ} \text{ total de nascidos vivos de mães residentes}} \times 1.000 \quad (1)$$

No Brasil, o Instituto Brasileiro de Geografia e Estatística (IBGE) é o responsável pela divulgação de informações e registros de estatísticas vitais (FRIAS *et al.*, 2013). Além do IBGE, o Sistema de Informações sobre Mortalidade do Ministério da Saúde (SIM/RS) apresenta-se como uma fonte alternativa de dados de óbitos no país. É por meio da implantação de um modelo de declaração de óbito (DO) padronizado nacionalmente que se torna possível atender às exigências legais e fornecer subsídios para a análise e interpretação das taxas de mortalidade infantil (MINISTÉRIO DA SAÚDE, 2017).

De modo geral, a mortalidade infantil é classificada em dois grupos distintos conforme a idade da criança (SZWARCOWALD *et al.*, 1997). O primeiro deles é a mortalidade neonatal, quando o óbito ocorre durante as primeiras quatro semanas de vida da criança e está diretamente relacionado às causas perinatais e anomalias congênitas, enquanto que o segundo é caracterizado como elemento pós neonatal ou

tardio e está associado à desnutrição ou contágio de doenças infecciosas (POLLARD, YUSUF e POLLARD., 1974).

Durante a década de 80 o desenvolvimento social brasileiro já era considerado muito inferior quando comparado ao dos países desenvolvidos, com destaque para as estatísticas de saúde e as precárias condições de vida da população. A desinformação e desigualdade social prejudicam, principalmente, direito à vida de suas crianças (ALMEIDA e SZWARCOWALD, 2012).

Neste contexto, estudos como o de Szwarcwald et al. (1997), apontam que desde muito antes da década de 90 há um aumento expressivo no número de mortes infantis no Brasil, sendo mais acentuado em regiões mais pobres (GARCIA e SANTANA, 2011). Algumas das causas apontadas são a necessidade de deslocamento intermunicipal para o parto, a precariedade dos serviços e a falta de integração com a saúde básica, que se apresentam como riscos crescentes para a mortalidade infantil (ALMEIDA e SZWARCOWALD, 2012).

Porém, esse aumento nem sempre se manteve constante. Houve um decréscimo na taxa de natalidade e no coeficiente de mortalidade infantil em todas as regiões do Brasil, no período de 2000 a 2009, sendo que a taxa de decréscimo anual no Nordeste, maior região do Brasil, foi de 6,0% (FRIAS *et al.*, 2013). Esse decréscimo no coeficiente de mortalidade infantil também é relatado no estudo de Nascimento et al. (2014), que aponta que o coeficiente era de 20,4 e diminuiu para 12,1 a cada 1.000 nascidos vivos.

Analisando o cenário internacional, Zhou e Bennett (2017) destacam que, na China, houve a redução do índice de mortalidade infantil em 6,5% durante o período de 1990 a 2015, porém que o país ainda apresenta um número significativo de óbitos infantis, atrás apenas da Índia, dentre os países analisados no estudo. A Índia, por sua vez, também apresenta um índice de mortalidade materna acentuado, com 281,8 mortes para cada 100 mil nascidos vivos.

Os índices de mortalidade infantil de um país refletem seu desenvolvimento econômico e social (SAHU *et al.*, 2015). Nesse contexto, Szwarcwald et al. (1997) apontam que os óbitos infantis decorrentes de doenças infecciosas são importantes indicadores da situação de pobreza, enquanto que a desnutrição, falta de saneamento básico e as deficiências de cuidados primários são causas da mortalidade infantil por diarreia. Além disso, têm-se a influência dos indicadores socioeconômicos que são

capazes de ampliar drasticamente os riscos de óbitos infantis em uma determinada população (MAKAROVA *et al.*, 2016).

Também deve-se levar em conta que a exposição materna à vários riscos diretos ou indiretos diminui as chances de sobrevivência da criança, uma vez que a saúde pré-natal e os cuidados obstétricos são afetados (ÖZDEMIR *et al.*, 2015). Assim, as mães que integram uma população sem acesso à água potável, eletricidade e saúde básica estão sujeitas aos riscos de mortalidade devido às precárias condições de saúde, compartilhando esses riscos com seus filhos (RAI *et al.*, 2017).

Ao analisar o âmbito social, o grande desafio é reduzir os óbitos infantis das classes desfavoráveis com intervenções estratégicas relacionadas à saúde (MOHAMMADI *et al.*, 2018). Para que essas estratégias sejam eficientes, é preciso conhecer, primeiramente as causas da mortalidade infantil para selecionar as estratégias mais eficientes na redução desses óbitos (RAI *et al.*, 2017). As causas da mortalidade infantil estão associadas com determinantes a nível individual, domiciliar e comunitário, além de indicar que as características físicas e sociais da vizinhança onde a criança vive podem afetar drasticamente as condições e os comportamentos ligados à sua saúde (ADEDINI *et al.*, 2015).

Evidências apontam que viver em um bairro com baixas condições econômicas e sociais aumenta significativamente as chances de mortalidade em crianças com menos de cinco anos de idade (ZYLBERSZTEJN *et al.*, 2018). Cabe ressaltar que a qualidade do sistema de saúde e dos fatores sociais variam de uma região geográfica para outra (O'HARE *et al.*, 2013), sendo que a omissão da informação de óbitos infantis ainda é um agravante (ALMEIDA e SZWARCOWALD, 2014).

Mesmo com a implementação de algumas estratégias para a melhorias da qualidade do atendimento a gestantes no Brasil, as ações que visam garantir o acesso igualitário à assistência ao parto ainda são insuficientes (ALMEIDA e SZWARCOWALD, 2012). Entretanto, o nível de escolaridade e a renda mensal materna, além do alcoolismo e tabagismo, são fatores que influenciam diretamente na mortalidade infantil, aumentando o risco de óbito (MOHAMMADI *et al.*, 2018).

Portanto, a melhora na qualidade dos sistemas de saúde, valorização dos dados nacionais e a divulgação de políticas de incentivo ao acompanhamento gestacional são fundamentais para a preservação da vida de crianças (ALMEIDA e SZWARCOWALD, 2014). Vale ressaltar que muitas das mortes podem ser evitáveis, principalmente aprimorando-se a organização, qualidade e o acesso aos serviços de

saúde (CARVALHO *et al.*, 2018). Logo, é importante identificar os óbitos que poderiam ter sido evitados por uma atenção adequada à saúde materno-infantil (NASCIMENTO *et al.*, 2014).

### 3 MATERIAIS E MÉTODOS

Esta fase da pesquisa consiste na elaboração do plano de trabalho que este estudo abrange, visando atingir os objetivos propostos e responder ao problema da pesquisa. Nesse contexto, no que se refere ao modo de realização da pesquisa a natureza é definida como aplicada, pois se utiliza da análise dos fundamentos já especificados (GIL, 2010). Em relação à categorização dos objetivos, a pesquisa é exploratória e descritiva por se tratar do levantamento de informações e aplicação de técnicas estatísticas específicas, explorando-se os efeitos gerados (LAKATOS e MARCONI, 2010).

A pesquisa também se caracterizou como descritiva porque permitiu identificar relações entre variáveis, coletando opiniões e observando fatos que possam influenciar no resultado final da pesquisa (MEDEIROS, 2013). Quanto ao tratamento dos dados, a forma de abordagem ao problema é quantitativa, pois foi mensurado dados numéricos. Além disso o estudo também é classificado como bibliográfico, pois se utilizou de conceitos publicados em livros e artigos científicos para estruturar e confirmar as informações descritas (GIL, 2010).

A revisão da literatura foi realizada com o apoio dos artigos científicos encontrados nas bases de dados: Web of Science (Capes), Plataforma SciELO e PubMed. Posteriormente foi realizada uma análise descritiva no estado do RS, das taxas de mortalidade infantil, estratificando por sexo (feminino e masculino), idade (Horas, varia de 01 a 23; Dias, varia de 01 a 29; Meses, varia de 01 a 11) raça/cor (branca, preta, amarela, parda e indígena), município de residência da mãe e local de ocorrência (hospital, outro estab. saúde, domicílio, via pública, outros e ignorado) além das informações referentes às mães: idade (idade da mãe em anos), escolaridade (1 a 3 anos de estudo, 4 a 7 anos, 8 a 11 anos, 12 e mais, ignorado e nenhuma) quantidade de filhos na gravidez (quantidade de filhos vivos) e, têm-se também as informações referentes ao parto (vaginal, cesáreo e ignorado).

Quanto à análise das séries temporais, foram utilizados os registros mensais relacionados às taxas de mortalidade infantil por município de residência da mãe, disponíveis no site do DATASUS – Departamento de Informática do SUS (<http://datasus.saude.gov.br/>), referente a janeiro de 2000 a junho de 2017, para o ajuste, e de julho a dezembro de 2017, para avaliação da qualidade dos modelos.

Foi realizada a modelagem ARIMA, seguindo os passos metodológicos propostos por Box-Jenkins (1976) e a modelagem  $\beta$ ARMA desenvolvida por Rocha e Cribari-Neto (2009), visando a obtenção do modelo de previsão mais eficiente, utilizando-se os softwares R v3.4.1 e *EViews v9.5 Student Version Lite*.

### 3.1 METODOLOGIA DE BOX-JENKINS

O método proposto por Box-Jenkins (1976) é utilizado com maior ênfase na realização de previsão em séries temporais, o qual também é conhecido como metodologia ARIMA. Este método destaca a análise das propriedades da própria série temporal, em que o  $Y_t$  é explicado pelos valores passados, defasados do próprio  $Y$ , o mesmo ocorre para o erro estocástico (GUJARATI e PORTER, 2011).

O modelo ARIMA consiste nas expressões identificadas como a ordem ( $p$ ) da parte autorregressiva (AR), com ordem de diferenciação ( $d$ ) e a ordem ( $q$ ) referente a médias móveis (MA), as quais caracterizam os modelos de Box-Jenkins denotados ARIMA ( $p, d, q$ ) (JERE, KASENSE e CHILYABANYAMA, 2017). Estes modelos são resultantes da combinação do componente AR, do filtro de Integração (I) e do componente MA, permitindo a modelagem da série por meio dos três componentes ou por meio de um subconjunto deles, resultando assim em outros modelos (MORETTIN e TOLOI, 2006).

Para Gujarati e Porter (2011) um modelo  $AR(p)$  é expresso pela Equação (2):

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + \varepsilon_t \quad (2)$$

Enquanto que o modelo  $MA(q)$  é dado pela Equação (3):

$$\tilde{Z}_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (3)$$

logo:  $\tilde{Z}_t = Z_t - \mu$

O parâmetro  $\theta_1$  explica a relação de  $\tilde{Z}_t$  com o  $\varepsilon_{t-1}$ , considerando  $i = 1, 2, \dots, q$ .

A sazonalidade é outra característica que pode ser adicionada aos modelos ARIMA e representa os picos que se repetem anualmente na série temporal (MONTGOMERY, JENNINGS e KULAHCI, 2015). Os modelos com componentes sazonais são denominados ARIMA sazonal ou SARIMA (*Seasonal Autoregressive*

*Integrated Moving Average*) e são denotados como SARIMA  $(p, d, q)(P, D, Q)_s$  sendo que o "s" representa a ordem da sazonalidade (MORETTIN e TOLOI, 2006). O modelo SARIMA genérico é expresso pela equação (4)

$$(1 - \phi_p B^p)(1 - \phi_p B^s)(1 - B^d)(1 - B^D)y_t = (1 - \theta_q B^q)(1 - \theta_Q B^s)\varepsilon_t \quad (4)$$

Em que a primeira parte da equação (4) representa a parte autorregressiva do modelo, sendo os parâmetros autorregressivos  $p$ , os parâmetros sazonais autorregressivos  $P$ , e as diferenças  $d$ , enquanto que as diferenças sazonais são representadas por  $D$  com período correspondente a  $s$ . A segunda parte da equação (4) indica, respectivamente, os componentes de médias móveis com ordem  $q$  e os parâmetros sazonais de médias móveis  $Q$  com  $s$  períodos (MAKRIDAKIS, WHEELWRIGHT e HYNDMAN., 2008).

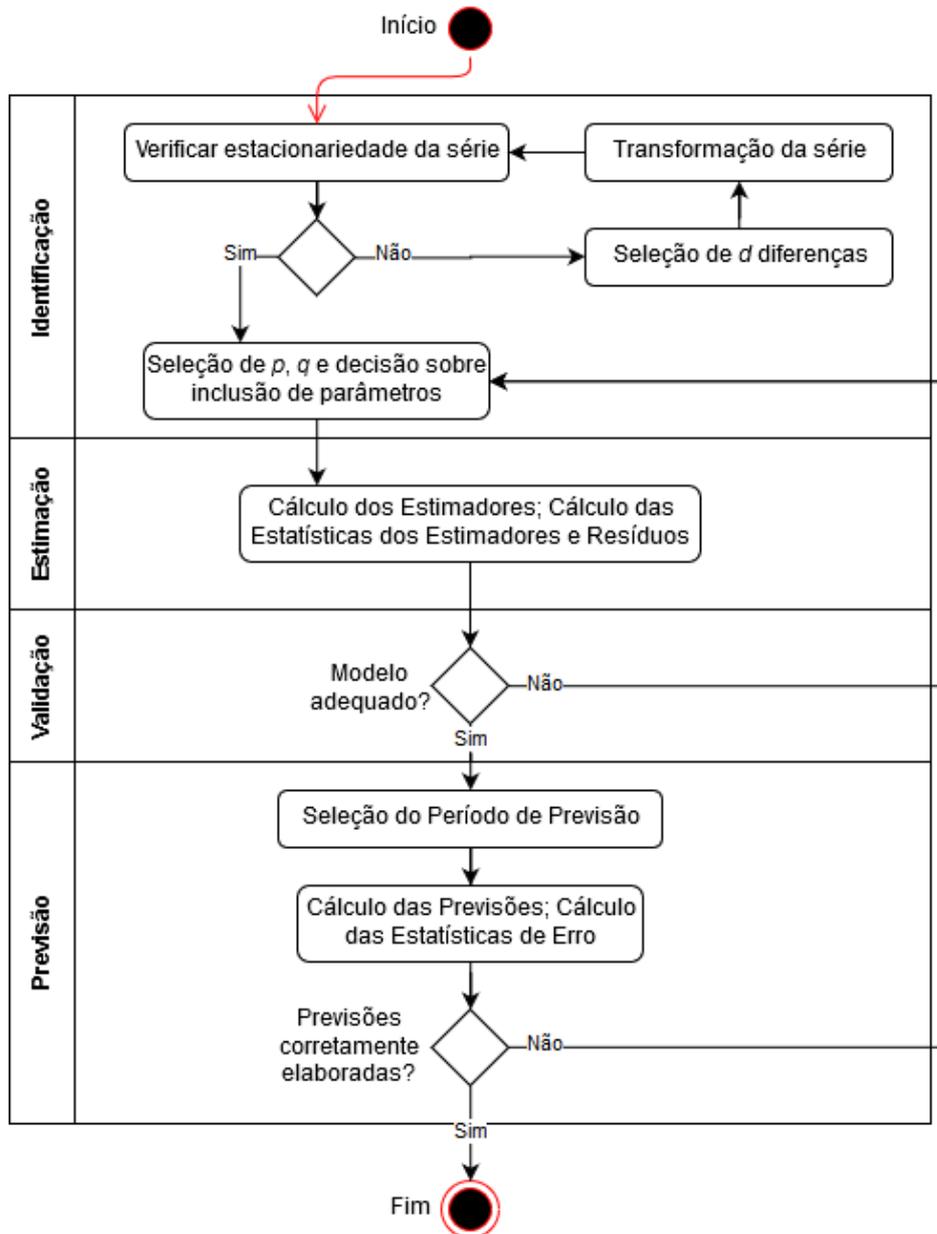
Contudo, para a utilização da metodologia ARIMA, primeiramente, é preciso garantir que haja significância dos parâmetros estimados ao longo do tempo, para isso, é necessário atender ao pressuposto básico da estacionariedade da série temporal (MONTGOMERY, JENNINGS e KULAHCI, 2015).

A série temporal é estacionária quando se desenvolve, ao longo do tempo um certo equilíbrio em torno da média, ou seja, as leis de probabilidade não mudam no processo, se mantendo constantes (SOUZA, 2016). Quando a série temporal não apresentar tal comportamento, podem ser aplicadas transformações, sendo que a ordem de integração (I) considerada, corresponde ao número de diferenças necessárias para converter a série não estacionária em estacionária (FAVA, 2000).

O método proposto por Box-Jenkins representa um ciclo interativo que indica o modelo predictor mais eficiente levando em conta as características de autocorrelação presentes em cada série temporal analisada (BOX e JENKINS, 1976; AZEVEDO e CAMPOS, 2016).

Na Figura 1 pode-se observar as etapas metodológicas do processo de modelagem proposto por Box-Jenkins (1976).

Figura 1 – Etapas metodológicas de Box-Jenkins



Fonte: adaptado de Pereira e Requeijo (2008)

Na primeira etapa, ao aplicar o modelo de Box-Jenkins a uma série de dados é preciso determinar quais os filtros AR, I e MA, compõe o processo gerador da série e quais são suas respectivas ordens. Esta escolha se dá por meio da condição de estacionariedade da série temporal. Realizada a verificação da estacionariedade e aplicada as devidas diferenciações, quando necessário, parte-se para a etapa de escolha do melhor modelo predictor, baseando-se na análise da Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP). Gujarati e Porter

(2011) explicam que os parâmetros da FAC e FACP na série temporal fornecem subsídios ao pesquisador para a escolha do(s) modelo(s) que melhor representa(m) o processo gerador da série temporal.

Posteriormente à etapa de identificação dos filtros  $(p, d, q)$  que representam a série temporal realiza-se a estimação dos parâmetros " $p$ " e " $q$ ". Tal estimação pode ser realizada por meio do método da maximização da Função de Verossimilhança condicional ou ainda, por meio da técnica dos Mínimos Quadrados Ordinários (BUENO, 2008; FAVA, 2000). É relevante destacar que na etapa de estimação dos parâmetros o modelo escolhido seja o mais parcimonioso possível (BOX e JENKINS, 1976). Para auxiliar na decisão do número de parâmetros que devem ser introduzidos ao modelo, utiliza-se o critério informacional de *Akaike* – AIC (AKAIKE, 1973) e o critério informacional *Bayesiano* – BIC (SCHWARZ, 1978), os quais são calculados por meios das equações (5) e (6) respectivamente.

$$AIC(p, q) = \ln\sigma_{\varepsilon}^2 + \frac{2(p + q)}{n} \quad (5)$$

$$BIC(p, q) = \ln\sigma_{\varepsilon}^2 + (p + q) \frac{\ln(n)}{n} \quad (6)$$

Onde  $p$  e  $q$  representam a ordem dos parâmetros identificados,  $n$  é o tamanho da amostra,  $\ln$  é o logaritmo neperiano, enquanto que  $\sigma_{\varepsilon}^2$  representa a variância estimada dos erros.

Os critérios AIC e BIC são denominados critérios penalizadores pois consideram o número de parâmetros do modelo, como também a variância dos erros gerados, sendo que quanto menores os valores de AIC e BIC mais adequado é o modelo ajustado (MORETTIN, 2011; AZEVEDO e CAMPOS, 2016). Tais critérios indicam a adequabilidade do modelo e foram considerados neste estudo para determinar, entre os modelos propostos aqueles que possuem o melhor desempenho.

Após a etapa de identificação dos modelos e estimação dos parâmetros, tem-se a validação que procura avaliar se o modelo ajustado é capaz de descrever o comportamento dos valores da série temporal analisada. No entendimento de Bueno (2008), os erros encontrados no modelo ajustado devem atender aos pressupostos de ruído branco, ou seja, média zero, independência e variância constante.

Ainda, considerou-se as medidas de acurácia como critérios alternativos para identificar o melhor ajuste entre os modelos concorrentes (MONTGOMERY, JENNINGS e KULAHCI, 2015). As medidas mais utilizadas estão ilustradas no Quadro 1.

Quadro 1 – Medidas de acurácia dos modelos estimados

<b>Sigla</b>	<b>Descrição</b>	<b>Equação</b>
MAE	<i>Mean Absolute Error</i> (erro médio absoluto)	$\frac{\sum_{t=1}^n  \varepsilon_t }{n}$
MAPE	<i>Mean Absolute Percentage Error</i> (erro percentual médio absoluto)	$\frac{\sum_{t=1}^n  \varepsilon_{t_t}/Z_t  * 100}{n}$
MSE	<i>Mean Squared Error</i> (erro quadrático médio)	$\frac{\sum_{t=1}^n (\varepsilon_t)^2}{n}$
RMSE	<i>Root Mean Squared Error</i> (raiz do erro quadrático médio)	$\frac{\sqrt{\sum_{t=1}^n (\varepsilon_t)^2}}{n}$

Fonte: adaptado de Bueno (2008)

Onde  $\varepsilon_t$  representa o erro de previsão no instante  $t$ , o  $n$  representa o número de observações enquanto que o  $Z_t$  corresponde ao valor real no tempo  $t$  e  $t = 1, 2, \dots, n$ .

Por fim, tem-se a previsão da série temporal, última etapa da metodologia proposta por Box-Jenkins (1976) que consiste em realizar previsões *in-sample*, calcular as estatísticas MAE, MAPE e RMSE para futuros valores da série, utilizando-se os modelos escolhidos.

Além do método de previsão proposto por Box-Jenkins foi utilizada a metodologia  $\beta$ ARMA em busca do modelo de previsão mais eficaz para a modelagem das séries correspondentes às taxas de mortalidade infantil no estado do RS, Brasil.

### 3.2 METODOLOGIA $\beta$ ARMA

A modelagem de dados distribuídos em um intervalo limitado segue, principalmente, duas abordagens: aplicar transformações nos dados para obter uma linha temporal e usar um dos modelos tradicionais de análise; ou modelar diretamente

os dados originais (CASARIN, DALLA VALLE e LEISEN, 2012). Essa segunda abordagem foi aprimorada por McKenzie (1985), com a introdução de um novo processo beta autorregressivo para séries temporais no intervalo unitário (0,1). Uma contribuição mais recente é apresentada por Ferrari e Cribari-Neto (2004) e Rocha e Cribari-Neto (2009), com modelos beta autorregressivos e a metodologia  $\beta$ ARMA, respectivamente.

Essa metodologia baseia-se na distribuição beta ( $\beta$ ), que é utilizada para modelar variáveis aleatórias de interesse distribuídas em um intervalo delimitado por dois valores escalares conhecidos  $a$  e  $b$ , onde  $a < b$ . Os escalares  $a$  e  $b$ , geralmente assumem, respectivamente, os valores 0 e 1, casos de modelagem de percentuais e taxas, por exemplo (GUPTA e NADARAJAH, 2004; PALM, BAYER e CINTRA, 2016). A função densidade de probabilidade da distribuição beta é dada pela equação (7), onde  $r > 0$ ,  $s > 0$ , e  $\Gamma(\cdot)$  é a função gama.

$$f(y; r, s) = \frac{\Gamma(r + s)}{\Gamma(r)\Gamma(s)} y^{r-1}(1 - y)^{s-1}, \quad 0 < y < 1, \quad (7)$$

A média  $\mathbb{E}$  e a variância  $var$  de  $y$  são dadas pelas equações (8).

$$\mathbb{E}(y) = \frac{r}{r + s} \quad var(y) = \frac{rs}{(r + s)^2(r + s + 1)} \quad (8)$$

Os modelos de regressão utilizados para analisar dados relacionados a outras variáveis, em geral, não são apropriados para avaliar situações em que a resposta está restrita ao intervalo unitário (0, 1), uma vez que eles podem gerar valores para a variável de interesse que excedem os limites inferiores e superiores, portanto Ferrari e Cribari-Neto (2004) propõem em seu trabalho um modelo de regressão para variáveis aleatórias cujo retorno é uma distribuição beta, a fim de atender essas situações em que a variável dependente ( $y$ ) está contida no intervalo unitário, ou seja,  $0 < y < 1$ .

O modelo beta autorregressivo de médias móveis ( $\beta$ ARMA), proposto por Rocha e Cribari-Neto (2009), parte do trabalho de Ferrari e Cribari-Neto (2004) para buscar a padronização de um modelo para séries temporais voltadas a taxas e proporções. O modelo  $\beta$ ARMA, assim como os modelos ARIMA, pressupõe a estacionariedade dos dados analisados (MARTINS JÚNIOR, 2014). Porém, a série de

interesse pode apresentar alguns picos típicos que retratam a ausência de estacionariedade, sendo que, nestes casos, ainda é considerado viável a utilização deste modelo mediante a aplicação de diferenças na série (ROCHA e CRIBARI-NETO, 2009).

Este modelo é largamente utilizado devido ao fato de que a função densidade de probabilidade pode ser crescente ou decrescente e unimodal, lembrando que sua utilização compreende os casos em que o resultado é restrito ao intervalo que compreende  $a$  e  $b$ , ou seja,  $a < b$ . Assim, aceita-se a transformação de  $(y - a)/(b - a)$ , para modelar a variável  $y_t$  em um intervalo unitário padrão. Neste caso, tem-se a densidade condicionada de  $y_t$  considerando  $F_{t-1}$ , expressa pela equação (9).

$$f(y_t|F_{t-1}) = \frac{\Gamma(\phi)}{\Gamma(\mu_t \phi)\Gamma((1 - \mu_t) \phi)} y_t^{\mu_t \phi - 1} (1 - y_t)^{(1 - \mu_t) \phi - 1}, 0 < y_t < 1, \quad (9)$$

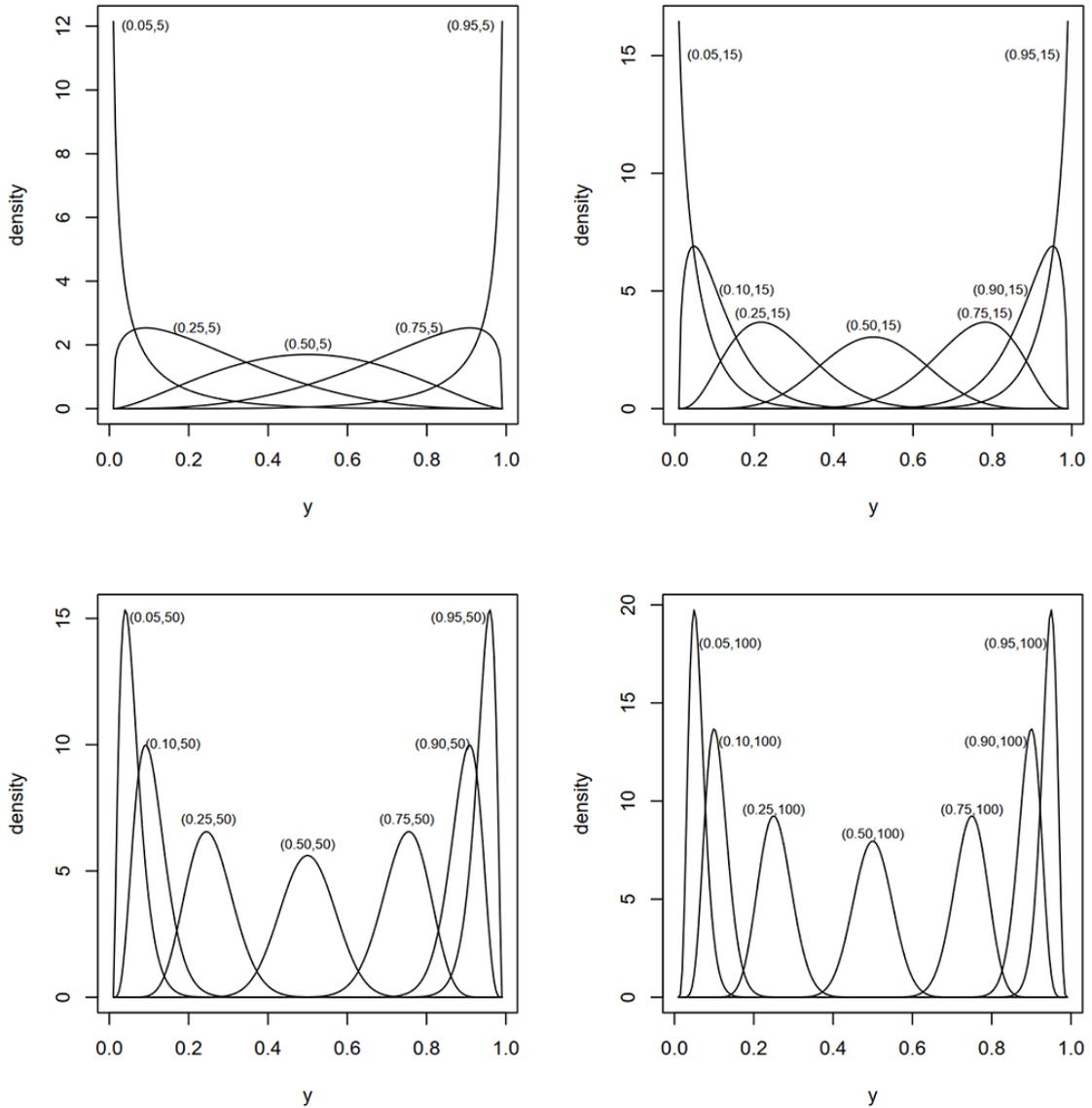
$$\text{Logo, } \mathbb{E}(y) = \mu \text{ e } \text{var}(y) = \frac{V(\mu)}{1 + \phi}$$

Neste caso,  $F_{t-1}$  é a menor  $\sigma$ -álgebra em que as variáveis  $y_1, y_2, \dots, y_{t-1}$  são mensuráveis, seguindo distribuição Beta  $(\mu_t, \phi)$  e  $V(\mu) = \mu(1 + \mu)$  é considerada uma função de variância em que  $\mu$  representa a média e  $\phi$  um parâmetro de precisão. Pressupõe-se que quando o parâmetro de  $\mu$  é fixado, o valor de  $\phi$  é inversamente proporcional a variância de  $y$ . Esta é considerada uma reparametrização da densidade beta original (ROCHA e CRIBARI-NETO, 2009).

Deste modo, tem-se  $y_t$ , considerando  $t = 1, \dots, n$ , o número de ocorrências da variável de interesse  $y$  no sentido em que  $y_t$  apresenta a densidade explicada pela equação (9) em que a média é dada por  $\mu_t$  e a precisão constante por  $\phi$ .

A densidade expressa pela equação (9) é bastante flexível para modelar proporções, uma vez que sua curva pode ter formatos diferentes, tais como U, sino, ou J (com cauda para a esquerda ou para a direita), dependendo do valor de dois parâmetros (BAYER e BAYER, 2015; PAJHEDE, 2017). Exemplos podem ser visualizadas na Figura 2, juntamente com seus valores de  $\mu$  e  $\phi$ .

Figura 2 – Densidades de Beta para diferentes combinações de  $(\mu, \phi)$



Fonte: Ferrari e Cribari-Neto (2004)

Rocha e Cribari-Neto (2009) salientam que o modelo  $\beta$ ARMA  $(r, s)$ , recomendado para modelar a esperança condicional de  $y$ , pode ser descrito por meio da equação (10):

$$g(\mu_t) = \alpha + \sum_{i=1}^r \phi_i g_{(y_{t-i})} + \sum_{j=1}^s \theta_j v_{t-j}, \quad (10)$$

Dado que  $\alpha \in \mathfrak{R}$ , é considerado constante,  $\phi_i$  são os parâmetros autorregressivos e  $\theta_j$  os parâmetros de médias móveis,  $v_t = y_t - \mu_t$  mede o erro da

equação de médias móveis e  $g(-)$  é descrito como uma função de ligação, podendo assumir duas diferenças  $g: (0,1) \rightarrow \mathbb{R}$ , como também ocorre no modelo de regressão beta. Seja nos modelos ARIMA ou nos parâmetros do modelo  $\beta$ ARMA, é relevante esclarecer que ambos são estimados via máxima verossimilhança (ROCHA e CRIBARI-NETO, 2009; BAYER e BAYER, 2015; NIELSEN, 2017).

#### 4 RESULTADOS E DISCUSSÕES

A caracterização das taxas de mortalidade infantil, registradas no estado do RS, está representada nessa seção por meio da análise descritiva e estratificação conforme Tabela 1. Ainda, nesse capítulo há a representação gráfica das séries originais demonstrando a prevalência das mortalidades infantis no estado do RS, Brasil.

Tabela 1 – Medidas descritivas das taxas de mortalidade infantil a cada 1.000 nascidos vivos no estado do RS, de 2000 a 2017

Estatística descritiva	Rio Grande do Sul		
	Taxas de mortalidade infantil de 2000 a 2017	Mortalidade infantil	Nascimentos
Média	12,70	153,4	12.090
Mediana	12,11	142,5	11.981
Desvio padrão	2,52	39,62	1.163,75
Coeficiente de Variação (%)	0,20	0,2582	0,0963
Mínimo	6,79	84	9.514
Máximo	20,01	278	15.915

Fonte: Elaborado pela autora

O maior número mensal de mortes infantis no RS no período de 2000 a 2017 ocorreu em julho de 2000, com 278 mortes contabilizadas. Nesse mês, a taxa de mortalidade infantil foi de 18,66 mortes a cada 1.000 nascidos vivos e houve 14.902 nascimentos, sendo que 146 (53%) eram do sexo masculino, 124 (45%) tinham idade entre 3 horas a 4 meses e 214 (77%) tinham pele branca. A capital do estado, Porto Alegre, foi o município de residência com o maior número de mortes infantis, 35 (13%), enquanto que a cidade de Pelotas ocupou a segunda posição com 17 (6%) casos. Quanto ao local de ocorrência do óbito, o hospital foi citado em 210 (76%) casos.

Em relação às informações das mães, em 57 (21%) ocorrências a faixa etária era de 19 a 32 anos, porém esse campo foi ignorado em 120 (43%) situações. Em 55 (20%) ocorrências essas mulheres apresentavam de 4 a 7 anos de estudo. A gravidez foi única (apenas um bebê) em 148 (53%) situações e a duração da gestação foi de 37 a 41 semanas em 65 (23%) ocorrências. Em 103 (37%) situações o parto foi

registrado como normal ou vaginal, enquanto que em 158 (57%) casos a mortalidade infantil ocorreu depois do parto e em 46 (17%) ocorrências a mãe tinha um filho vivo.

Em relação à taxa mensal de mortalidade infantil no estado do RS, a maior taxa registrada para o período de 2000 a 2017 ocorreu em junho de 2002 com 20,01 mortes a cada 1.000 nascidos vivos (245 mortes infantis registradas dentre 12.246 nascimentos). Dentre essas mortes, 140 (57%) eram do sexo masculino e 116 (47%) possuíam idade entre 01 dia e 5 meses, porém há 129 (53%) casos em que a idade das vítimas foi ignorada. Quanto à raça ou cor, 194 (79%) eram brancos. O hospital foi o principal local de ocorrência do óbito infantil em 201 (82%) situações.

No tocante às características das mães, 64 (26%) apresentavam idade entre 19 e 28 anos, porém em 181 (74%) casos a idade não foi informada. Quanto à escolaridade, 85 (35%) apresentaram 4 a 7 anos de estudo. Em 167 (68%) situações a gravidez foi única e em 69 (28%) a gestação teve duração de 37 a 41 semanas. O parto foi registrado como vaginal ou normal em 125 (51%) óbitos. O óbito infantil ocorreu depois do parto em 182 (74%) casos, sendo que em 49 (20%) ocorrências a mãe não tinha nenhum filho vivo.

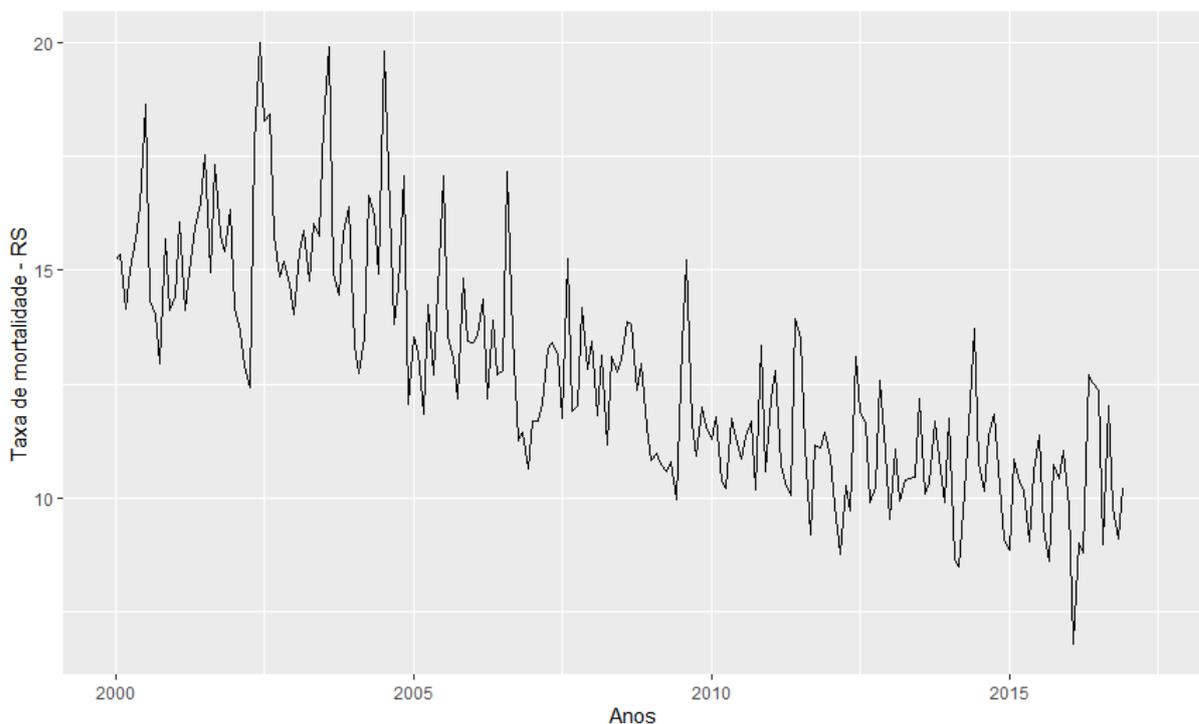
A menor taxa mensal de mortalidade infantil no RS ocorreu em fevereiro de 2016 com 6,790 mortes a cada 1.000 nascidos vivos (84 mortes infantis registradas dentre 12.371 nascimentos). Esse mês também representou o menor número de mortes infantis no período de 2000 a 2017. Destas 84 mortes, 53 (63%) eram do sexo masculino, 72 (86%) brancos e em 34 (40%) situações a vítima apresentava idade entre 3 horas e 2 meses. Em 80 (95%) casos o local de ocorrência do óbito infantil foi o hospital. Quanto ao município de residência, Porto Alegre foi citado em 13 (15%) casos, seguido de Novo Hamburgo com 7 (8%) casos. A idade da mãe foi ignorada em 50 (60%) ocorrências e em 34 (40%) situações elas apresentavam de 8 a 11 anos de estudo. Em 24 (29%) casos já apresentavam um filho vivo, e em 71 (85%) situações a gravidez foi única. A duração da gravidez foi de 37 a 41 semanas em 26 (31%) ocorrências e em 45 (54%) ocorreu o parto cesáreo. A morte da criança ocorreu depois do parto em 79 (94%) casos.

Para a análise estratificada por município (com base no município de residência da mãe), foi realizada a soma de todas as mortes, assim como a soma de todos os nascimentos no período de 2000 a 2017. Neste caso, o município de Cacique Doble apresentou a maior taxa de mortalidade infantil para o estado do RS, de 28,19 mortes a cada 1.000 nascidos vivos (1.135 nascidos e 32 mortes), seguido por Alto Alegre

com 27,78 mortes a cada 1.000 nascidos vivos (288 nascidos e 8 mortes) e São Valério do Sul com 26,75 mortes a cada 1.000 nascidos vivos (785 nascimentos e 21 mortes). A capital Porto Alegre apresentou uma taxa de 11,3 mortes a cada 1.000 nascidos vivos (348.053 nascidos e 3.932 mortes), enquanto que a média do estado é 12,73 mortes a cada 1.000 nascidos vivos.

A Figura 3 representa a série original das taxas mensais de mortalidade infantil a cada 1.000 nascidos vivos para o estado do RS, para uma análise gráfica a fim de melhor compreender o comportamento das variáveis analisadas. É possível observar o comportamento das taxas de mortalidade infantil durante os 210 meses considerados. Destaca-se que a série não apresentou evidências de estacionariedade.

Figura 3 – Comportamento das taxas mensais de mortalidade infantil a cada 1.000 nascidos vivos no estado do RS, no período de 2000 a 2017



Fonte: Elaborado pela autora

Posteriormente à análise gráfica da série original, tem-se a realização da primeira etapa da metodologia de Box-Jenkins, a análise das funções de autocorrelação (FAC) e autocorrelação parcial (FACP). Além disso, realizou-se os testes ADF e KPSS com o intuito de verificar a condição de estacionariedade e

identificar os possíveis modelos para a série temporal analisada. Os valores dos testes de raiz unitária ADF e KPSS, para o estado do RS podem ser observados Tabela 2.

Tabela 2 – Principais valores para os testes de raiz unitária ADF e KPSS para a série em nível

Estado	Série em nível		
	ADF*	p-valor	KPSS**
RS	-2,036	0,563	3,745

MacKinnon (1996). \*\*Kwiatkowski et al. (1992)

Na série de mortalidade infantil analisada, a hipótese de existência de raiz unitária não foi rejeitada, visto que ambos os testes convergem seus resultados, desse modo optou-se pela diferença em primeira ordem da série, tornando-a estacionária.

Determinar as ordens  $p$  e  $q$  dos modelos é uma etapa fundamental no ajuste dos modelos de séries temporais ARIMA e  $\beta$ ARMA (PALM, BAYER e CINTRA, 2016). Nessa perspectiva, a seleção do modelo considerou uma busca computacional exaustiva com o intuito de minimizar o AIC (AKAIKE, 1973) e o BIC (SCHWARZ, 1978). O espaço da busca foi restringido aos modelos com ordens menores ou iguais a 6, ou seja, foi considerado  $0 \leq p \leq 6$  e  $0 \leq q \leq 6$ .

A utilização do modelo  $\beta$ ARMA para a previsão das taxas de mortalidade infantil destacou-se, uma vez que a distribuição Beta é restrita ao intervalo de possíveis valores para as taxas de mortalidade, ou seja, o intervalo (0,1). Assim, valores preditos menores que 0% ou maiores 100% para as taxas de mortalidade não são possíveis de ocorrer. Porém, ao se considerar os tradicionais modelos da classe ARIMA, utilizados para modelagem de taxas, podem ocorrer valores irreais da taxa de mortalidade, acima de 100% ou abaixo de 0%, uma vez que o suporte da distribuição normal é toda a reta real (BAYER e BAYER, 2015).

Pelo fato de as taxas de mortalidade infantil a cada 1.000 nascidos vivos não estarem restritas a este intervalo unitário, utilizou-se neste estudo a transformação sugerida por Cribari-Neto e Zeileis (2009) a fim de permitir o uso do  $\beta$ ARMA nos dados analisados.

Na sequência, tem-se a Tabela 3 na qual pode ser observada a análise de diagnóstico dos modelos SARIMA e  $\beta$ ARMA selecionados para o estado do RS, Brasil.

Tabela 3 – Modelos SARIMA e  $\beta$ ARMA ajustados para as taxas de mortalidade infantil a cada 1.000 nascidos vivos no estado do RS, Brasil

Estado	Modelo	Parâmetros	p-valor	AIC	BIC
RS	$SARIMA(1,1,1) (1,01)_{12}$	$\Phi_1 = 0,180$ $\theta_1 = -0,923$ $\Phi_{12} = 0,923$ $\theta_{12} = 0,762$	0,017 < 0,001 < 0,001 < 0,001	741,71	758,42
	$\beta ARMA(4,6)$	$\Phi_4 = 0,282$ $\theta_6 = 0,119$	0,0000 0,0246	-129,6316	-116,2432

Fonte: Elaborado pela autora

Foram ajustados quatro modelos concorrentes para a série de dados. Na Tabela 3, foram utilizados o AIC e o BIC para identificar aquele com melhor ajuste à série de taxas de mortalidade infantil no estado do RS, Brasil. Sendo o modelo para a série do estado do RS foi  $\beta ARMA(4,6)$  pois apresentou o menor AIC e BIC.

A qualidade do ajuste da modelagem  $\beta ARMA$  também se mostrou superior no estudo de Palm e Bayer (2018) que analisou as taxas de energia armazenada no Sul do Brasil, no período 2009 a 2015. Tal conclusão também foi obtida no estudo de Bayer, Tondolo e Müller (2018) ao analisar um banco de dados relacionado a um processo de fabricação de pneus, bem como nas análises dos dados de umidade relativa do ar nas cidades de Brasília (BAYER, TONDOLO e MÜLLER, 2018), Manaus (PUMI *et al.*, 2019) e Santa Maria (BAYER, CINTRA e CRIBARI-NETO, 2018). Porém não há estudos que utilizaram tal modelagem para o ajuste das taxas de mortalidade infantil.

Imediatamente após, verificou-se que os modelos ajustados atenderam a todos os pressupostos para a caracterização de ruído branco. Assim, posteriormente a seleção dos modelos realizou-se a última etapa da metodologia de Box-Jenkins, a previsão, que neste estudo foi efetuada para seis instantes de tempo posteriores à observação 204 (referente a junho de 2017).

Logo, os valores originais e previstos por ambas as metodologias podem ser observados na Tabela 4 e na Figura 4

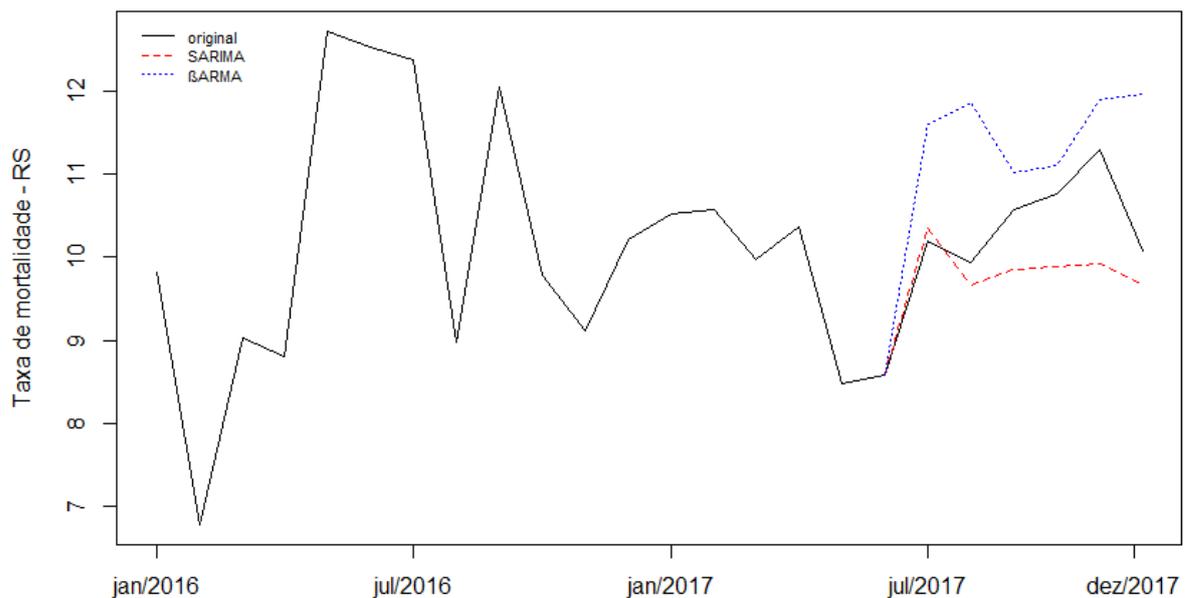
Tabela 4 – Resultado para as previsões das taxas de mortalidade no estado do RS, Brasil, com base na previsão *out-sample* para os meses de julho a dezembro de 2017

Estado	Período	Série Original	Previsão SARIMA	Previsão $\beta$ ARMA
RS	Jul/2017	10,186335	10,341512	11,60439
	Ago/2017	9,943425	9,661384	11,86188
	Set/2017	10,573882	9,859713	11,01921
	Out/2017	10,760808	9,877287	11,10782
	Nov/2017	11,287988	9,916187	11,89055
	Dez/2017	10,070404	9,685453	11,95435

Fonte: Elaborado pela autora

Na Figura 4 pode-se observar as previsões, bem como os valores reais para o período. Observa-se que as previsões do modelo  $\beta$ ARMA, no período de julho a dezembro de 2017 na região Sul do Brasil apresentaram melhor ajuste quando comparadas ao modelo SARIMA.

Figura 4 – Comparação da série original com a previsão SARIMA e  $\beta$ ARMA para as taxas de mortalidade infantil a cada 1.000 nascidos vivos no estado do RS, Brasil



Fonte: Elaborado pela autora

Pode-se observar na Tabela 4 que o modelo  $\beta$ ARMA previu valores mais próximos aos dados originais em comparação com a metodologia ARIMA. Com o objetivo de verificar a importância de considerar estimadores não viesados, realizou-se a comparação dos modelos em relação aos valores reais e previstos. Essa comparação foi realizada por meio da análise da Raiz Quadrada do Erro-Médio (RMSE), do Erro Médio Absoluto (MAE) e do Erro Percentual Médio Absoluto (MAPE). Essas medidas avaliam a diferença entre os valores originais e os valores previstos e podem ser observadas na Tabela 5.

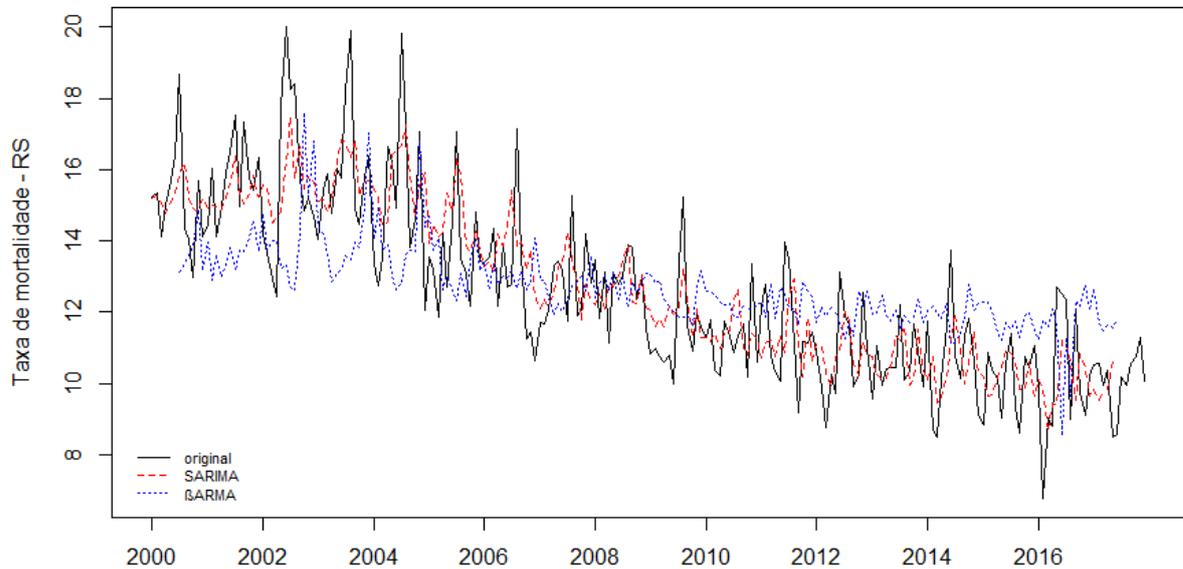
Tabela 5 – Medidas de acurácia para as previsões das taxas de mortalidade infantil no estado do RS, Brasil, para os meses de julho a dezembro de 2017

RMSE			MAE			MAPE		
SARIMA	$\beta$ ARMA	%	SARIMA	$\beta$ ARMA	%	SARIMA	$\beta$ ARMA	%
1,371	1,286	-6%	1,093	1,103	+1%	8,861	10,783	+22%

Fonte: Elaborado pela autora

Porém, na Tabela 5 é perceptível por meio dos valores percentuais (%) que os valores atribuídos às medidas de acurácia RMSE, MAE e MAPE não são menores (exceto para o RMSE) quando se considera os valores previstos para o modelo  $\beta$ ARMA. Na Figura 5 observa-se as previsões dentro da amostra. A metodologia  $\beta$ ARMA apresentou melhor ajuste uma vez que exhibe os menores valores de AIC e BIC.

Figura 5 – Valores originais e previstos dentro da amostra para as taxas de mortalidade infantil a cada 1.000 nascidos vivos no estado do RS, Brasil



Fonte: Elaborado pela autora

Após o ajuste dos melhores modelos de previsão para as taxas de mortalidade infantil no estado do RS, Brasil, foi realizada uma análise nos resíduos a qual contou com a aplicação do teste de Ljung-Box (1978), que tem como objetivo verificar a existência de autocorrelação nos resíduos. Esse teste indicou que os resíduos dos modelos não apresentaram autocorrelação, evidenciando bom ajuste dos modelos. Com relação ao teste de normalidade, realizou-se o teste de Shapiro-Wilk (1965) que apresentou indicativo de normalidade dos resíduos para os modelos ajustados.

Por fim, vale ressaltar que a saúde dos indivíduos depende principalmente das condições de saúde do seu local de residência, as desigualdades sociais existentes e as condições de vida precárias. Além do mais, no Brasil o aumento da renda está diretamente associado às melhores condições de saúde, isso significa que, em geral, os óbitos infantis estão concentrados em regiões mais pobres com acentuada desnutrição infantil, reflexos da fragilidade nas condições de vida, saneamento básico e saúde pública.

Esse fato, além de retratar as condições de assistência à saúde, está relacionado a diversas privações, principalmente em relação à educação, justificando a baixa escolaridade materna, além do aumento na exposição às doenças, resistência imunológica prejudicada e busca inadequada por serviços de saúde. Tais questões reforçam as reflexões acerca das condições de saúde neonatal e pós-neonatal que

resultam na mortalidade infantil, evidenciando a necessidade de priorização no planejamento de intervenções e estratégias direcionadas à saúde infantil.

## 5 CONCLUSÃO

O coeficiente de mortalidade infantil foi indicado por município de residência da mãe, além das características de vulnerabilidade dos bebês, das mães e do parto, para o estado do RS. Desse modo, este estudo apresentou uma análise comparativa da aplicação dos modelos ARIMA e  $\beta$ ARMA para a previsão das taxas de mortalidade infantil no estado do RS, Brasil. Trata-se da primeira aplicação para examinar a capacidade de predição destes modelos de séries temporais para as taxas de mortalidade infantil no estado do RS, Brasil.

A modelagem  $\beta$ ARMA assume a distribuição beta para a variável analisada. Os parâmetros deste modelo são estimados por meio da maximização da função de probabilidade condicional. Essas inferências apresentam bons resultados quando o tamanho da amostra é grande ( $> 30$ ), mas podem levar a distorções inferenciais no caso de amostras pequenas ( $< 30$ ).

Quando comparados os modelos SARIMA e  $\beta$ ARMA observou-se um melhor desempenho no ajuste do modelo  $\beta$ ARMA, no horizonte de tempo de 6 meses, visto que o melhor modelo foi  $\beta$ ARMA(4,6) para a série do estado do Rio Grande do Sul (RS). Este modelo se apresentou superior pois obteve o menor AIC e BIC. Nesse sentido o resultado condiz com o esperado visto que o  $\beta$ ARMA é um modelo específico de séries temporais contínuas.

Além disso, a análise dos resíduos dos modelos escolhidos também confirma a superioridade da metodologia  $\beta$ ARMA em descrever o comportamento das taxas de mortalidade infantil no estado do RS, Brasil. Neste contexto, considerando o fato de que há pouquíssimas aplicações do  $\beta$ ARMA, os resultados levantados apontam para um indício empírico de que a aplicação de tal modelo se mostrou positiva tanto em termos de qualidade do ajuste quanto aos resultados para previsões de séries temporais do tipo proporções ou taxas com intervalos limitados.

Dessa forma, considerando o crescente número de pesquisas utilizando-se a metodologia ARIMA (Jere, Kasense e Chilyabanyama (2017) Azevedo e Campos (2016), Aye et al. (2015)) entre outros, e  $\beta$ ARMA (Bayer e Bayer (2015), Casarin, Calla Valle e Leisen (2012), Ferrari e Cribari-Neto (2004)), entre outros, sugere-se, primeiramente, a aplicação dessas metodologias na análise da mortalidade infantil nos outros estados do Brasil. Além disso, é altamente encorajado a realização de estudos

comparando as metodologias para a previsão das taxas de mortalidade fetal, que possuem impactos sociais e humanos, tal qual a mortalidade infantil.

Por fim, conclui-se que as metodologias propostas podem contribuir de forma significativa para orientar no planejamento de políticas preventivas e educativas voltadas ao risco de um nascido vivo morrer durante o seu primeiro ano de vida.

## REFERÊNCIAS

- ADEDINI, S. A. et al. Regional variations in infant and child mortality in Nigeria: a multilevel analysis. **Journal of biosocial science**, v. 47, n. 2, p. 165-187, 2015.
- AKAIKE, H. Information theory and an extension of the maximum likelihood principle. **2nd International Symposium on Information Theory**, Tsahkadsor, Armenia, USSR, 1973.
- ALMEIDA, W. D. S. D.; SZWARCOWALD, C. L. Mortalidade infantil e acesso geográfico ao parto nos municípios brasileiros. **Revista de Saúde Pública**, v. 46, p. 68-76, 2012.
- ALMEIDA, W. D. S. D.; SZWARCOWALD, C. L. Mortalidade infantil nos municípios brasileiros: uma proposta de método de estimação. **Revista Brasileira de Saúde Materno Infantil**, v. 14, n. 4, p. 331-342, 2014.
- AYE, G. C. et al. Forecasting aggregate retail sales: The case of South Africa. **International Journal of Production Economics**, v. 160, p. 66-79, 2015.
- AZEVEDO, V. G.; CAMPOS, L. M. Combination of forecasts for the price of crude oil on the spot market. **International Journal of Production Research**, v. 54, n. 17, p. 5219-5235, 2016.
- BAYER, D. M.; BAYER, F. M. Previsão da Umidade Relativado Ar de Brasília por meio do Modelo Beta Autorregressivo de Médias Móveis. **Revista Brasileira de Meteorologia**, v. 30, n. 3, p. 319 - 326, 2015.
- BAYER, F. M.; CINTRA, R. J.; CRIBARI-NETO, F. Beta seasonal autoregressive moving average models. **Journal of Statistical Computation and Simulation**, v. 88, n. 15, p. 2961-2981, jul. 2018.
- BAYER, F. M.; TONDOLO, C. M.; MÜLLER, F. M. Beta regression control chart for monitoring fractions and proportions. **Computers & Industrial Engineering**, v. 119, p. 416-426, maio 2018.
- BOX, G. E. P.; JENKINS, G. M. **Time series analysis forecasting and control**. Oakland: Holden-Day, 1976. ISBN 0-8162-1104-3.
- BUENO, R. D. L. D. S. **Econometria de séries temporais**. São Paulo: Cengage Learning, 2008.
- CARVALHO, I. C. B. D. M. et al. Adaptation and validation of the World Health Organization's Safe Childbirth Checklist for the Brazilian context. **Revista Brasileira de Saúde Materno Infantil**, Recife, v. 18, n. 2, p. 401-418, jun. 2018.
- CASARIN, R.; DALLA VALLE, L.; LEISEN, F. Bayesian model selection for beta autoregressive processes. **Bayesian Analysis**, v. 7, n. 2, p. 385-410, 2012.
- CRIBARI-NETO, F.; ZEILEIS, A. **Beta Regression in R**. WU Vienna University of Economics and Business. Vienna. 2009. (98).

- FAVA, V. L. **Testes de raízes unitárias e co-integração. Manual de econometria: nível intermediário.** São Paulo: Atlas, 2000.
- FERRARI, S.; CRIBARI-NETO, F. Beta Regression for Modelling Rates and Proportions. **Journal of Applied Statistics**, 31, n. 7, 2004. 799-815.
- FRIAS, P. G. D. et al. Correção de informações vitais: estimação da mortalidade infantil, Brasil, 2000-2009. **Revista de Saúde Pública**, v. 47, p. 1048-1058, 2013.
- GARCIA, L. P.; SANTANA, L. R. Evolução das desigualdades socioeconômicas na mortalidade infantil no Brasil, 1993-2008. **Ciência & Saúde Coletiva**, v. 16, p. 3717-3728, 2011.
- GIL, A. C. **Como Elaborar Projetos de Pesquisa.** 5. ed. São Paulo: Atlas, 2010.
- GUJARATI, D. N.; PORTER, D. C. **Econometria Básica.** 5. ed. Porto Alegre: AMGH Editora Ltda, 2011.
- GUPTA, A. K.; NADARAJAH, S. **Handbook of Beta Distribution and Its Applications.** [S.I.]: CRC Press, 2004. ISBN 9780824753962.
- JERE, S.; KASENSE, B.; CHILYABANYAMA, O. Forecasting Foreign Direct Investment to Zambia: A Time Series Analysis. **Open Journal of Statistics**, v. 7, p. 122-131, 28 fev. 2017.
- KWIATKOWSKI, D. et al. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? **Journal of econometrics**, v. 54, n. 1-3, p. 159-178, 1992.
- LAKATOS, E. M.; MARCONI, M. D. A. **Fundamentos de Metodologia Científica.** 7. ed. São Paulo: Atlas, 2010.
- LJUNG, G. M.; BOX, G. E. P. On a measure of lack of fit in time series models. **Biometrika**, v. 65, n. 2, p. 297-303, 1978.
- MACKINNON, J. G. Numerical distribution functions for unit root and cointegration tests. **Journal of applied econometrics**, v. 11, n. 6, p. 601-618, 1996.
- MAKAROVA, N. et al. Comparative analysis of premature mortality among urban immigrants in Bremen, Germany: a retrospective register-based linkage study. **BMJ open**, v. 6, n. 3, p. e007875, 2016.
- MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. **Forecasting methods and applications.** [S.I.]: John Wiley & Sons, 2008.
- MARSHALL JUNIOR, I. et al. **Gestão da qualidade e processos.** Rio de Janeiro: FGV, 2012.
- MARTINS JÚNIOR, H. **O uso da distribuição beta T generalizada nos erros do modelo Garch : uma aplicação no preço do café.** Universidade Federal de Lavras - UFLA. Lavras - MG, p. 101. 2014.
- MCKENZIE, E. An autoregressive process for beta random variables. **Management Science**, v. 31, n. 8, p. 988-997, 1985.

MEDEIROS, J. B. **Redação Científica**: a prática de fichamentos, resumos, resenhas. 11. ed. São Paulo: Atlas, 2013.

MINISTÉRIO DA SAÚDE. Portal da Saúde. **Ministério da Saúde**, 2017. Disponível em: <<http://portalsaude.saude.gov.br>>. Acesso em: 17 jul. 2017.

MOHAMMADI, Y. et al. Estimating Child Mortality Rate and its Trend in Hamadan Province, Western Iran from 1990 to 2016: Implications for Sustainable Development Goal. **Journal of research in health sciences**, v. 18, p. 1-4, 2018.

MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. **Introduction to time series analysis and forecasting**. [S.l.]: John Wiley & Sons, 2015.

MORETTIN, P. A. **Econometria financeira - Um curso em séries temporais financeiras**. 2. ed. São Paulo: Blucher, 2011.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. 3. ed. São Paulo: Egard Blücher, 2006.

NASCIMENTO, S. G. D. et al. Mortalidade infantil por causas evitáveis em uma cidade do Nordeste do Brasil. **Revista Brasileira de Enfermagem**, v. 67, n. 2, 2014.

NAVARRO, A. C. et al. Decisões de investimento e rentabilidade futura: Estudo empírico com companhias abertas não financeiras. **Contabilidade, Gestão e Governança**, Brasília, v. 16, n. 1, p. 19-34, abr. 2013. ISSN 1984-3925.

NIELSEN, T. P. **Tales From the Unit Interval: Backtesting, Forecasting and Modeling**. Department of Economics, University of Copenhagen. Copenhagen. 2017.

O'HARE, B. et al. Income and child mortality in developing countries: a systematic review and meta-analysis. **Journal of the Royal Society of Medicine**, v. 106, n. 10, p. 408-414, 2013.

ÖZDEMİR, R. et al. Özdemir, Raziye, et al. "Reliable mortality statistics for Turkey: Are we there yet?" **BMC public health**, v. 15, p. 1-23, 2015.

PAJHEDE, T. **A conditionally beta distributed time-series model with application to monthly US corporate default rates**. University of Copenhagen. Copenhagen, p. 31. 2017.

PALM, B. G.; BAYER, F. M. Bootstrap-based inferential improvements in beta autoregressive moving average model. **COMMUNICATIONS IN STATISTICS-SIMULATION AND COMPUTATION**, v. 47, p. 977-996, 2018.

PALM, B. G.; BAYER, F. M.; CINTRA, R. J. **Intervalos de Predição no Modelo Beta Autorregressivo de Médias Móveis**. Universidade Federal de Santa Maria - UFSM. Santa Maria, RS, p. 1-102. 2016.

PEREIRA, Z. L.; REQUEIJO, J. G. **Qualidade**: Planejamento e Controle Estatístico de Processos. Lisboa: FCT/UNL, 2008. ISBN 9789898022653.

POLLARD, A. H.; YUSUF, F.; POLLARD, G. N. **Demographic techniques**. Sydney, Austrália: [s.n.], 1974.

- PUMI, G. et al. Beta autoregressive fractionally integrated moving average models. **Journal of Statistical Planning and Inference**, v. 200, p. 196–212, 2019.
- RAI, S. K. et al. Causes of and contributors to infant mortality in a rural community of North India: evidence from verbal and social autopsy. **BMJ open**, v. 7, n. 8, p. e012856, 2017.
- ROCHA, A. V.; CRIBARI-NETO, F. Beta autoregressive moving average models. **Test**, v. 18, n. 3, p. 529-545, 2009.
- SAHU, D. et al. Levels, trends & predictors of infant & child mortality among Scheduled Tribes in rural India. **The Indian journal of medical research**, v. 141, n. 5, p. 709, 2015.
- SCHWARZ, G. E. Estimating the dimension of a model. **Annals of Statistics**, v. 6, p. 461-464, 1978.
- SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). **Biometrika**, v. 52, n. 3/4, p. 591-611, 1965.
- SOUZA, F. M. **Modelos De Previsão: aplicações à energia elétrica - ARIMA- ARCH- AI e ACP**. Curitiba: Appris, 2016. ISBN 9788547302481.
- SZWARCWALD, C. L. et al. Mortalidade infantil no Brasil: Belíndia ou Bulgária? **Cadernos de Saúde Pública**, v. 13, p. 503-516, 1997.
- ZHOU, L.; BENNETT, S. **Asian Health Benchmarking Technical Report for Waitemata DHB and Auckland DHB**. Auckland. 2017.
- ZYLBERSZTEJN, A. et al. Child mortality in England compared with Sweden: a birth cohort study. **The Lancet**, v. 391, n. 10134, p. 2008–2018, 2018.