

Francisco Canindé Assis de Oliveira

Análise de Quebra Estrutural no Número de Casos de Covid-19 no Rio Grande do Norte

Natal - RN

17 de janeiro de 2022

Francisco Canindé Assis de Oliveira

Análise de Quebra Estrutural no Número de Casos de Covid-19 no Rio Grande do Norte

Monografia de Graduação apresentada ao Departamento de Estatística do Centro de Ciências Exatas e da Terra da Universidade Federal do Rio Grande do Norte como requisito parcial para a obtenção do grau de Bacharel em Estatística.

Universidade Federal do Rio Grande do Norte

Centro de Ciências Exatas e da Terra

Departamento de Estatística

Orientador: Prof. Dr. Marcus Alexandre Nunes

Natal - RN

17 de janeiro de 2022

Universidade Federal do Rio Grande do Norte - UFRN
Sistema de Bibliotecas - SISBI
Catalogação de Publicação na Fonte. UFRN - Biblioteca Setorial Prof. Ronaldo Xavier de Arruda - CCET

Oliveira, Francisco Canindé Assis de.

Análise de quebra estrutural no número de casos de Covid-19 no Rio Grande do Norte / Francisco Canindé Assis de Oliveira. - 2022.

28f.: il.

Monografia (Bacharelado) - Universidade Federal do Rio Grande do Norte, Centro de Ciências Exatas e da Terra, Departamento de Estatística. Natal, 2022.

Orientador: Prof. Dr. Marcus Alexandre Nunes.

1. Estatística - Monografia. 2. COVID-19 - Monografia. 3. Pandemia - Monografia. 4. Quebras estruturais - Monografia. 5. Séries temporais - Monografia. I. Nunes, Marcus Alexandre. II. Título.

RN/UF/CCET

CDU 519.2



**MINISTÉRIO DA EDUCAÇÃO E DO ESPORTE
UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS EXATAS E DA TERRA
CURSO DE GRADUAÇÃO EM ESTATÍSTICA**

**ATA DE DEFESA DE MONOGRAFIA
BACHARELADO EM ESTATÍSTICA**

Às 15h00min do dia 31 de janeiro de 2022, por meio da ferramenta de transmissão Google Meet, compareceu para defesa da monografia do curso de graduação em Estatística o aluno *Francisco Canindé Assis De Oliveira* tendo como Título da Monografia “Análise de Quebra Estrutural no Número de Casos de Covid-19 no Rio Grande do Norte”. Constituíram a Banca Examinadora os professores Marcus Alexandre Nunes (presidente), Luz Milena Zea Fernández (examinadora) e Mariana Correia de Araújo (examinadora). Após a apresentação e as observações dos membros da banca avaliadora, ficou definido que o trabalho foi considerado aprovado.

Prof. Dr. Marcus Alexandre Nunes (DEST/UFRN)
Orientador

Profa. Dra. Luz Milena Zea Fernández (DEST/UFRN)
Examinador

Profa. Dra. Mariana Correia de Araújo (DEST/UFRN)
Examinador



Emitido em 31/01/2022

ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO Nº 1/2022 - CCE/CCET (12.25)

(Nº do Protocolo: NÃO PROTOCOLADO)

(Assinado digitalmente em 02/02/2022 17:29)
ADRIANA MARCIA NASCIMENTO DE OLIVEIRA
AUXILIAR EM ADMINISTRACAO
ADM/CCET (12.01)
Matrícula: 1110117

(Assinado digitalmente em 02/02/2022 18:59)
LUZ MILENA ZEA FERNANDEZ
PROFESSOR DO MAGISTERIO SUPERIOR
EST/CCET (12.02)
Matrícula: 2083593

(Assinado digitalmente em 02/02/2022 17:46)
MARCUS ALEXANDRE NUNES
PROFESSOR DO MAGISTERIO SUPERIOR
EST/CCET (12.02)
Matrícula: 1066308

(Assinado digitalmente em 02/02/2022 17:46)
MARIANA CORREIA DE ARAUJO
PROFESSOR DO MAGISTERIO SUPERIOR
EST/CCET (12.02)
Matrícula: 2312009

Para verificar a autenticidade deste documento entre em <https://sipac.ufrn.br/documentos/> informando seu número: **1**
, ano: **2022**, tipo: **ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO**, data de emissão: **02/02**
/2022 e o código de verificação: **2b2ead9a8b**

À minha mãe Dona Telma.

Agradecimentos

Agradeço primeiramente a Deus sem ele não sou nada, a minha mãe Dona Telma por seu apoio por sempre estar comigo por puxar minhas orelhas quando se é preciso, à minha irmã Francisca Maria por sempre me motivar a continuar e não desistir, ao meu irmão Francisco de Assis pelo companheirismo e amizade e ao meu irmão mais velho José Horlando.

A meus amigos de infância Lucas Rafael, Ewerton Rodrigues, Gleison Américo e Isaías Silva, por me aturarem como amigo e por sempre me darem apoio nas minhas decisões.

A Luiz Antônio do Nascimento, que além de ser um ótimo professor um grande amigo, sempre disposto a ouvir minhas reclamações e desabafos e me aconselhar nas minhas decisões.

A amigos que fiz na residência de graduação da UFRN, Marcos Luan, Caio Matheus, Tardelli Jeferson, Luíz Gustavo, Rene Airton, Kevin Santos, José Veraldo e Valmir Correia são muitos, mas esses quero citar, pois melhoram a minha estadia neste ilustre local.

A amizades que surgiram durante meu estágio no Projeto Governo Cidadão sendo elas, Thyago Fernandes, Lucas Moura Galego, Lucas Felipe, Mavigson Silva, Danielle Cavalcanti e Monica Cruz que compartilharam maravilhosos momentos e conhecimentos.

Ao Professor Marcus Alexandre Nunes por aceitar me orientar neste trabalho e por ter me acompanhado durante minha graduação, as professoras Luz Milena Zea Fernández e Mariana Correia de Araújo por aceitarem o convite para compor a banca examinadora e aos demais professores do departamento de estatística por compartilharem seus conhecimentos comigo.

“É preciso cuidado com a obesidade mental. Ou seja, colocar muito de informação na mente e nada em prática.”

Caio Carneiro

Resumo

A pandemia da COVID-19 trouxe consigo um risco iminente de colapso dos sistemas de saúde, com impactos negativos na taxa de mortalidade, direta e indiretamente. O presente estudo tem por objetivo investigar se existem diferentes ondas de contaminação pela doença no estado do Rio Grande do Norte. Mais especificamente, pretende-se estimar a partir do método proposto por Bai e Perron (1998) a possível existência de quebras estruturais na série e relacioná-las aos acontecimentos do período de 26/02/2020 a 21/06/2021. Para tanto, foram utilizados os dados do número de infectados pelo COVID-19 no Rio Grande do Norte, obtidos através do site do Laboratório de Inovação Tecnológica em Saúde (LAIS). Os resultados apontaram quatro quebras estruturais i) a primeira por volta do dia 24 de maio de 2020 ii) a segunda no dia 11 de setembro de 2020 iii) a terceira no dia 15 de janeiro de 2021 iv) a quarta no dia 10 de abril de 2021. Tais evidências sugerem que os efeitos do lockdown observados pelos decretos estaduais, efeitos das festas de final de ano e eventos não permitidos durante o período de carnaval podem ter relação com o número de infectados pela COVID-19 no Rio Grande do Norte.

Palavras-chave: COVID-19. Pandemia. Quebras Estruturais. Séries Temporais.

Abstract

The COVID-19 pandemic brought with it an imminent risk of the collapse of health systems, with negative impacts on the mortality rate, both directly and indirectly. The present study aims to investigate whether there are different waves of contamination by the disease in the state of Rio Grande do Norte. More specifically, it is intended to estimate from the method proposed by Bai and Perron (1998) the possible existence of structural breaks in the series and relate them to the events of the period from 02/26/2020 to 06/21/2021. For this purpose, data on the number of people infected by COVID-19 in Rio Grande do Norte were used, obtained through the website of the Laboratory of Technological Innovation in Health (LAIS). The results showed four structural breaks i) the first around May 24, 2020 ii) the second on September 11, 2020 iii) the third on January 15, 2021 iv) the fourth on April 10, 2021. Such evidence suggests that the effects of the lockdown observed by state decrees, effects of the end-of-year parties and events not allowed during the carnival period may be related to the number of infected by COVID-19 in Rio Grande do Norte.

Keywords: COVID-19. Pandemic. Structural breaks. Time Series.

Lista de ilustrações

Figura 3.1 – Série do número de casos confirmados da COVID-19 no Rio Grande do Norte.	18
Figura 4.1 – Série do número de casos confirmados da COVID-19 no Rio Grande do Norte com estimativas das médias dos intervalos de confiança.	24

Lista de tabelas

Tabela 4.1 – Estatísticas descritivas da série Número de casos confirmados de COVID-19 no Rio Grande do Norte	22
Tabela 4.2 – Valores das estatísticas SQR e CIB para diferentes estimativas de quebras estruturais m	22
Tabela 4.3 – Datas das quebras identificadas pelo procedimento de Bai e Perron (1998)	23
Tabela 4.4 – Intervalos de confiança para às datas das quebras identificadas pelo procedimento de Bai e Perron (1998).	23

Sumário

	Lista de tabelas	11
1	INTRODUÇÃO	13
1.1	Objetivo	14
1.2	Justificativa	14
1.3	Estrutura do Trabalho	14
2	REVISÃO BIBLIOGRÁFICA	15
2.1	Uma Visão Sobre Pandemias	15
2.2	Quebras Estruturais	15
3	METODOLOGIA	17
3.1	Processos Estocásticos e Séries Temporais Estacionárias.	17
3.2	Método de Procura por Quebras Estruturais	18
3.2.1	Teste de m Quebras versus $m + 1$ Quebras	20
3.3	Crítério de Informação Bayesiano	20
4	APLICAÇÃO	22
5	CONSIDERAÇÕES FINAIS	26
	REFERÊNCIAS	27

1 Introdução

Entre o final de 2019 e início de 2020, o mundo começava a entrar em um status de pandemia, onde uma doença ou enfermidade espalha-se por vários ou todos os lugares do planeta, tal pandemia causada pelo vírus SARS-CoV-2, mais conhecido como coronavírus ou COVID-19. Segundo Farias (2020), o novo coronavírus foi descoberto em 31/12/2019, com seus primeiros casos registrados em Wuhan, na China. Em janeiro de 2020 foram surgindo casos em outros países, sendo os primeiros casos confirmados no Brasil foram registrados em fevereiro do mesmo ano, como mostrado em Cavalcante et al. (2020), mais especificamente em 26 de fevereiro de 2020. Em 12 de março de 2020 foi registrado o primeiro caso no Rio Grande do Norte, como apresentado em Gonçalves et al. (2020).

Cavalcante et al. (2020) relata que até 16 de maio de 2020 o número de casos da COVID-19 já ultrapassava a faixa de 4 milhões com mais de 300 mil óbitos, sendo as Américas e a Europa os continentes mais atingidos. Ainda conforme o estudo de Cavalcante et al. (2020), no período de 31 de dezembro de 2019 a 16 maio de 2020, os países que apresentavam os maiores números de infectados eram os Estados Unidos da América, com mais de 1.4 milhões de casos, seguido da Rússia, com mais de 262 mil e Reino Unido, com mais de 236 mil. Nesta época, o Brasil encontrava-se com o quarto maior número de infectados com aproximadamente 233 mil casos.

Segundo a Organização Pan-Americana de Saúde (OPAS) apud Gonçalves et al. (2020), até o dia 4 de outubro de 2020 foram confirmados no Brasil mais de 4 milhões de casos, sendo que o Rio Grande do Norte apresentava mais de 70 mil infectados.

Em meados de 2020 e com o avanço da pandemia, estudos relacionados ao SARS-CoV-2 foram tornando-se mais frequentes, tais como “A pandemia COVID-19”, de Ciotti et al. (2020), que fala sobre o surgimento do vírus, o desenvolvimento de vacinas, o impacto da pandemia e sobre uma possível segunda onda, que chegou a ocorrer plenamente em todos os lugares afetados por ela, “Educação e a pandemia COVID-19”, de Daniel (2020), que fala sobre a situação da educação durante a pandemia, a mudança para a forma de ensino remoto, como as instituições se saíam perante ao pouco tempo disponível para adaptação na forma de ensino e a relação com os pais e alunos e “Análise de Sobrevivência dos Infectados pela COVID-19 no Estado do Rio Grande do Norte”, de Cavalcanti (2021), que faz uso do método de Kaplan-Meier e do modelo Cox de risco para investigar os determinantes da probabilidade e do tempo de sobrevivência dos infectados pela COVID-19 no estado do Rio Grande do Norte.

Este trabalho faz a análise de quebra estrutural na série do número de casos confirmados de COVID-19 no Rio Grande do Norte. Dessa maneira, segundo Shikida,

Paiva e Junior (2016), entende-se por quebra estrutural uma ou mais mudanças no nível da série, na dispersão e/ou inclinação. No entanto, existem diversos métodos que procuram pela existência e a estimação de quebras estruturais, que serão apresentados no Capítulo 2. Contudo, o método proposto por Bai e Perron (1998) será utilizado, por ser o mais atual e mostrar resultados significativos em trabalhos referenciados na revisão bibliográfica.

1.1 Objetivo

Este trabalho tem como base os dados do número de infectados pelo coronavírus no Rio Grande do Norte obtidos através do site do Laboratório de Inovação Tecnológica em Saúde (LAIS) no período de 26/02/2020 a 21/06/2021. Os dados mostravam que a variação da série oscila bastante, logo a ideia de fazer um estudo referente a ela. Sendo assim, o presente estudo tem como objetivo averiguar a possível existência de quebras estruturais na série do número de casos confirmados de COVID-19 no Rio Grande do Norte no período considerado. Caso estas quebras sejam detectadas, desejamos estudar e fazer a associação com suas possíveis causas.

1.2 Justificativa

Em decorrência da relevância sobre estudos relacionados ao SARS-CoV-2 desde o seu descobrimento até os dias atuais, devido a escassez e a importância de estudos em relação a séries temporais e quebras estruturais em dados da área da saúde e a relevância do tema que é a COVID-19, este trabalho ajudará no entendimento dos acontecimentos que possam ter servido de causas para tais quebras estruturais na série em questão.

1.3 Estrutura do Trabalho

No Capítulo 2 realizamos uma revisão bibliográfica na qual encontra-se uma explicação sobre pandemias, alguns exemplos de surtos pandêmicos que já ocorreram no planeta e o levantamento do referencial teórico sobre o método de quebras estruturais proposto por Bai e Perron (1998). No Capítulo 3 tratamos do detalhamento do procedimento utilizado e do critério de escolha para o modelo mais adequado. A aplicação do método em dados reais e os resultados obtidos são mostrados no Capítulo 4. No Capítulo 5 são encontradas as considerações finais deste trabalho.

2 Revisão Bibliográfica

2.1 Uma Visão Sobre Pandemias

Analisando a história da humanidade, nos deparamos com diversos momentos em que situações relacionadas à saúde tomaram proporções relevantes, chegando ao status do que conhecemos hoje como pandemia.

No momento atual, a Organização Mundial da Saúde (OMS) conceitua pandemia como a “disseminação mundial de uma nova doença e o termo passa a ser usado quando uma epidemia, surto que afeta uma região, se espalha por diferentes continentes com transmissão sustentada de pessoa para pessoa” (disponível em: FIOCRUZ (2021), acesso em: 17/08/2021), ou ainda, conforme Freitas (2020), pandemia é o estágio mais grave de uma escala de incidência e pode ser decretada em casos onde uma patologia se espalha dentro de quadros epidêmicos por diversos países e continentes (disponível em: Freitas (2020), acesso em: 15/08/2021).

Existiram diversos momentos em que pandemias tomaram conta do cenário mundial, como exemplos podemos citar: a peste bubônica (1347-1353) que durante o período dizimou um quarto da população total da Europa, a gripe espanhola (1918-1920), que infectou 500 milhões de pessoas e estima-se que tenha matado entre 17-50 milhões, sendo umas das pandemias mais mortais da história da humanidade (disponível em: FIOCRUZ (2021), acesso em: 17/08/2021). Atualmente, a SARS-CoV-2, mais conhecida como COVID-19, que surgiu no final do ano de 2019, em Wuhan na China, (disponível em: Freitas (2020), acesso em: 15/08/2021), que deixou até o presente momento mais 381 milhões de infectados e o seu período de duração já passa de dois anos, (disponível em: Times (2022), acesso em 02/02/2022).

2.2 Quebras Estruturais

Estudos de quebras estruturais normalmente estão associados à econometria. Utilizados em dados de séries temporais, buscam encontrar mudanças nos níveis ou no mecanismo gerador das séries temporais. Alguns estudos buscam fazer inferências sobre acontecimentos históricos que possam ser motivos de tais mudanças, como por exemplo Shikida, Paiva e Junior (2016).

Este trabalho fez uso do método de detectar quebras estruturais proposto por Bai e Perron (1998) para analisar a existência de quebras estruturais na série do preço do boi gordo no estado de São Paulo no período de 1954-2012. Foram levadas em consideração

a importância da bovinocultura na agropecuária brasileira e as características que sua série pode apresentar. Os resultados obtidos por Shikida, Paiva e Junior (2016) foram que, segundo os autores, intervenções governamentais, tais como desestruturação do mercado bovino devido ao Plano Cruzado e a estabilização de preços após o Plano Real, influenciaram em mudanças significativas no comportamento da série do preço do boi gordo.

Abreu (2018) analisa a trajetória e o processo de desenvolvimento da economia brasileira com ênfase no panorama sócio-histórico na industrialização do país. Os métodos de quebras estruturais utilizados foram os métodos propostos por Lanne, Lütkepohl e Saikkonen (2002a), Lanne, Lütkepohl e Saikkonen (2002b) e Bai e Perron (1998). Os resultados apontaram que a série do grau de abertura econômica apresentava quebra estrutural em 1993. Segundo a autora, na década de 1990 o país passava por um processo inflacionário de mais de 80% ao mês. No mesmo ano houve a proposta do programa econômico que incluía reformas centrais na abertura comercial, redução do papel do Estado como produtor de bens e serviços e na reorganização da administração pública federal.

Brady (2008) relacionou a suavização do consumo com as mudanças estruturais no crédito ao consumidor. Endo e Gomes (2018) sugeriram que os métodos usados por Brady (2008) não foram adequados. Sendo assim, este fato levou Endo e Gomes (2018) a fazerem uso das metodologias de quebras estruturais propostas por Bai e Perron (1998) e Hall, Han e Boldea (2012) para tornar sua informação verídica. Os testes apontaram que a conclusão feita por Brady (2008) de que as mudanças estruturais tinham influências na suavização do consumo é errônea. Sendo assim, os resultados obtidos por Endo e Gomes (2018) levaram à conclusão de que a renda não mostrou-se estatisticamente significativa para explicar o consumo.

Durante o levantamento bibliográfico realizado para o desenvolvimento e escrita desta monografia, foram encontrados poucos trabalhos sobre séries temporais e análise de quebras estruturais em dados da saúde, tais como o trabalho de Kale, Fernandes e Nobre (2004), que faz uma análise sobre o padrão dos óbitos e internações, no período de 1995 a 1998, relacionados a diarreias em crianças menores de cinco anos de idade e o artigo de Cavalcanti (2021), que faz análise de sobrevivência em infectados pela COVID-19 no Rio Grande do Norte.

3 Metodologia

Análise e estudo de quebras estruturais normalmente estão relacionadas à econometria. Utiliza-se dados de séries temporais em busca de encontrar mudanças no nível da série. Neste trabalho é feita a análise na série do número de caso da COVID-19 no Rio Grande do Norte, com o objetivo de determinar se a série apresentou mudanças e se foram influenciadas por algum acontecimento.

Este capítulo, será dividido em três tópicos, sendo eles: uma breve definição de processo estocástico e séries temporais estacionárias, o método de quebras estruturais proposto por Bai e Perron (1998) e a definição do Critério de Informação Bayesiano.

3.1 Processos Estocásticos e Séries Temporais Estacionárias.

Um processo estocástico pode ser definido como um conjunto de variáveis aleatórias indexadas a uma variável, geralmente esta variável é o tempo conforme explicado em Alves e Delgado (1997a). Segundo Woodridge (2012), apud Signor (2020) uma definição de série temporal é um conjunto de observações ordenadas no tempo, ou seja, uma série temporal é a realização de um processo estocástico (disponível em: Educacional (2020), acesso em 03/02/2022). Sendo assim, uma série temporal pode ser dita como um processo estocástico.

Na econometria, as séries de tempo são geralmente não-estacionárias. Sendo assim, possuem média e variância que evoluem a medida que o tempo da série avança, ou seja, com raiz unitária. Nesse contexto, para verificar estacionariedade são feitos testes tais como o teste de Dickey-Fuller e o Dickey-Fuller Aumentado, como demonstrado em Abreu (2018).

Um processo estocástico é dito estacionário quando suas características não variam ao longo do tempo, Alves e Delgado (1997b) mostram que a função de distribuição das variáveis aleatórias que definem um processo estocástico não varia no tempo. Assim, uma série temporal é dita estacionária se a sua média e variância são constantes e sua auto correlação depende somente da defasagem de um tempo para o outro e não do período. As oscilações em períodos da série do número de casos da COVID-19 no Rio Grande do Norte apresentada a parti da Figura 3.1, fazem com que a série apresente média e variância que mudam conforme o tempo avança. Sendo assim, podemos presumir que a série é não-estacionária.

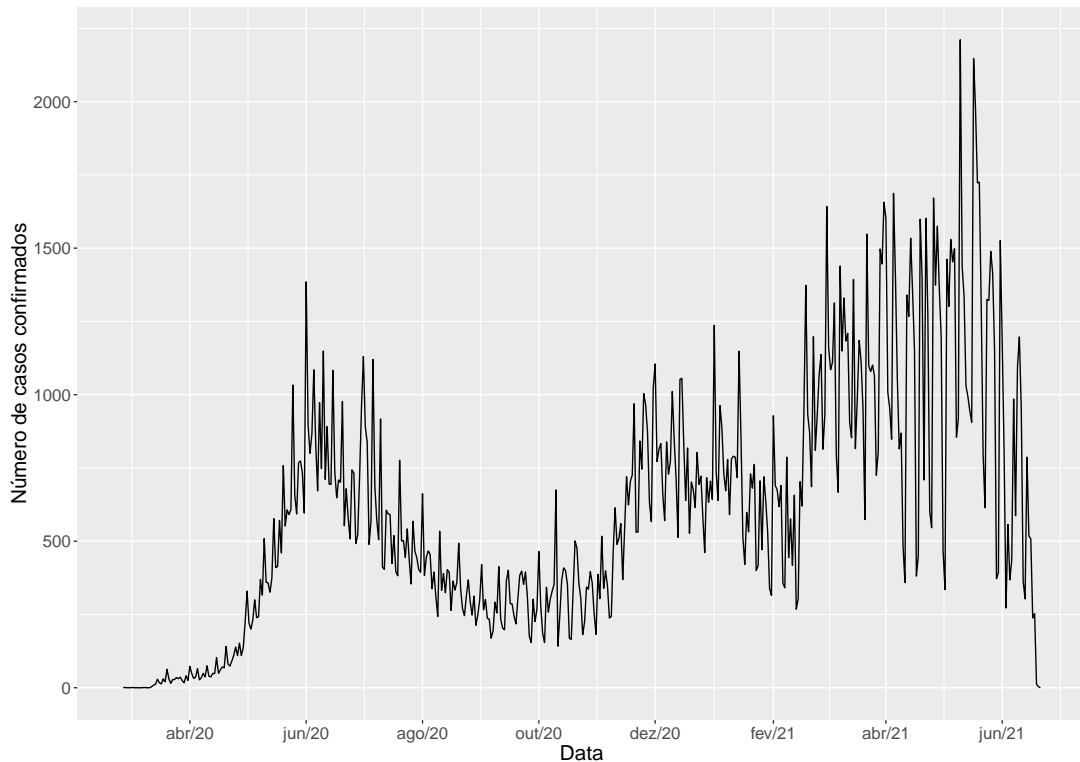


Figura 3.1 – Série do número de casos confirmados da COVID-19 no Rio Grande do Norte.

3.2 Método de Procura por Quebras Estruturais

Na literatura existem diferentes métodos para a procura de quebras estruturais, como já citados no Capítulo 2. Entretanto o método de quebras estruturais proposto por Bai e Perron (1998) além de ser o mais atual, mostrou-se ser mais eficaz nos resultados dos trabalhos apresentados na Revisão Bibliográfica.

O método de múltiplas quebras proposto por Bai e Perron (1998) trabalha em cima da hipótese de que exista pelo menos uma quebra, ou seja,

H_0 : Não há existência de quebra na série

H_1 : Existe pelo menos uma quebra na série

Assim, operando de forma sequencial, primeiramente testa se há existência dessa uma única quebra. Caso a hipótese nula de que não haja quebra na série for rejeitada, divide-se a amostra em duas (na data da quebra) e uma nova rodada de testes é realizada em cada sub-amostra. A sequência de testes continua até que a hipótese nula não seja rejeitada.

De acordo com Zeileis et al. (2003) apud Shikida, Paiva e Junior (2016), foi o trabalho de Bai (1994) que estabeleceu os primeiros fundamentos para os testes de

estimação de quebra, seguidos de trabalhos como podemos citar Bai (1997a), Bai (1997b) apud Zeileis et al. (2003) e Bai e Perron (1998).

O teste utilizado para identificar o número de quebras, também é usado para estimar a data da quebra, dividindo a amostra em sub-amostras em suas candidatas a quebras, e nas sub-amostras são estimados os parâmetros por Mínimos Quadrados Ordinários (*MQO*), com o cálculo e armazenamento da Soma dos Quadrados dos Resíduos (*SQR*). Em seguida, o teste busca as datas das quebras que minimizam o *SQR* total, de toda amostra. Entretanto, em alguns casos devemos escolher um modelo que seja mais simples possível e que explique bem os nossos dados o que será mostrado na Seção 3.3.

Bai e Perron (1998) desenvolveram um teste de quebras estruturais que leva em consideração a estimação de múltiplas quebras e modelos estimados por mínimos quadrados ordinários. Considere um modelo linear contido por múltiplas quebras dado por:

$$y_t = x_t' \beta + z_t' \delta_j + \epsilon_t. \quad (3.1)$$

em que $t = T_{j-1} + 1, \dots, T_j$, $j = 1, \dots, m + 1$, $T_0 = 0$ e $T_{m+1} = T$. No modelo apresentado, y_t é a variável dependente, x_t e z_t com dimensão $(p \times 1)$ e $(q \times 1)$, respectivamente, são vetores de covariáveis, β e δ_j ($j = 1, \dots, m+1$) vetores dos coeficientes e ϵ_t o termo de erro. Os pontos de quebras são dados pelos índices (T_1, \dots, T_m) , que são explicitamente desconhecidos. Assim, busca-se estimar os vetores dos coeficientes em conjunto dos pontos de quebras com as t observações de (y_t, x_t, z_t) . Percebe-se que o vetor β não é sujeito a mudanças, ou seja, somente δ_t sofre as mudanças das quebras.

Baseado nos princípios de mínimos quadrados, a soma dos quadrados dos resíduos é dada por:

$$SQR = \sum_{i=1}^{m+1} \sum_{t=T_{i-1}+1}^{T_i} (y_t - x_t' \beta - z_t' \delta_i)^2. \quad (3.2)$$

Minimizando a Equação 3.2, obtém-se as estimativas de β e δ_j , para cada uma das m partições (T_1, \dots, T_m) , que será denotada por $\{T_j\}$. Assim, as estimativas obtidas são $\hat{\beta}(\{T_j\})$ e $\hat{\delta}(\{T_j\})$.

Denotando a função *SQR* por $S_T(T_1, \dots, T_m)$ e fazendo a substituição dos parâmetros pelas suas estimativas, temos as estimativas dos pontos das quebras $(\hat{T}_1, \dots, \hat{T}_m)$, dadas por;

$$(\hat{T}_1, \dots, \hat{T}_m) = \operatorname{argmin}_{T_1, \dots, T_m} S_T(T_1, \dots, T_m). \quad (3.3)$$

Em que a minimização é feita em todas as m partições de (T_1, \dots, T_m) , tais que $T_i - T_{i-1} \geq 1$. Portanto, os parâmetros da regressão tornam-se as estimativas de mínimos quadrados

das m partições ($\{T_j\}$), ou seja, $\hat{\beta} = \hat{\beta}(\{T_j\})$ e $\hat{\delta}_t = \hat{\delta}_t(\{T_j\})$. Adaptado de Endo e Gomes (2018).

3.2.1 Teste de m Quebras versus $m + 1$ Quebras

Neste contexto são assumidas as hipóteses nula de m quebras versus $m + 1$ quebras. Em um modelo com m quebras utiliza-se a minimização global da soma dos quadrados dos resíduos para obter-se as datas das quebras estimadas, $\hat{T}_1, \dots, \hat{T}_m$. Sendo assim, testa-se a presença de uma quebra para cada um dos $m + 1$ segmentos que foram obtidos com as partições estimadas, $\hat{T}_1, \dots, \hat{T}_m$.

O objetivo é fazer $m + 1$ testes da hipótese nula de que haja quebra versus a hipótese de uma única quebra, fazendo uma vez em cada uma das $m + 1$ partições. O modelo com m quebras é rejeitado quando o menor valor do SQR calculado nas $m + 1$ partições for suficientemente menor que o SQR do modelo com m quebras. Selecionando a quebra associada com esse mínimo, então, o teste é dado por:

$$F(l + 1|l) = \frac{S_T(\hat{T}_1, \dots, \hat{T}_m) - \min_{1 \leq i \leq l+1} \inf_{T \in \Lambda_{i,n}} S_t(\hat{T}_1, \dots, T_{i-1}, \tau, \hat{T}_i, \dots, \hat{T}_l)}{\hat{\sigma}^2}. \quad (3.4)$$

Em que $\Lambda_{i,n} = \{\tau; \hat{T}_i + (\hat{T}_i - T_{i-1})n \leq \tau \leq \hat{T}_i - (\hat{T}_i - T_{i-1})n\}$ e $\hat{\sigma}^2$ é uma estimativa consistente de σ^2 (variância do modelo) sob a hipótese nula. Para mais detalhes consultar Endo e Gomes (2018) e Bai e Perron (1998).

3.3 Critério de Informação Bayesiano

Quando se fala em estimar um modelo, sabe-se que o modelo estimado não é verdadeiro, trata-se de uma aproximação da realidade. Em busca de chegar o mais próximo possível do conjunto de dados estudado e com uma melhor estimativa, é normal a escolha de um modelo que seja mais simples e que melhor explique o conjunto de dados em questão. Para isso existem alguns métodos que auxiliam nessa escolha entre modelos, tais como o Critério de informação de Akaike (CIA) e o Critério de Informação Bayesiano (CIB).

Proposto por Schwarz (1978) apud Emiliano et al. (2010), o Critério de Informação Bayesiano é dado por:

$$CIB = -2 \log f(y_t, x_t, z_t | \beta, \delta_j) + p \log t. \quad (3.5)$$

Em que $f(y_t, x_t, z_t | \beta, \delta)$ é o modelo mostrado em 3.1, p é o número de parâmetros a serem estimados e t é o número de observações da amostra (adaptados de Emiliano et al. (2010)).

Portanto, o modelo que deve ser escolhido será aquele com o número de quebras m que corresponda ao menor valor do Critério de Informação Bayesiano.

4 Aplicação

Neste capítulo serão apresentados os resultados da aplicação do método de quebras estruturais de Bai e Perron (1998) na série do número de infectados da COVID-19 no Rio Grande do Norte no período de 26 de fevereiro de 2020 até 21 de junho de 2021.

Antes de entrar na aplicação do teste de quebras estruturais proposto por Bai e Perron (1998), serão apresentadas algumas estatísticas descritivas da série do número de casos confirmados de COVID-19 no Rio Grande do Norte.

Tabela 4.1 – Estatísticas descritivas da série Número de casos confirmados de COVID-19 no Rio Grande do Norte

Média	Mediana	Variância	Desvio Padrão
600.3	531.0	181188.7	425.66

Primeiramente é feito um teste para detectar a existência de quebras na série. Para isso foi utilizada a função `breakpoints` do pacote `strucchang` no software R Development Core Team (2009). Resumidamente, a função `breakpoints` faz o cálculo dos pontos de quebras ideais, os procedimentos estão voltados a testar ou avaliar mudanças na estabilidade em modelos de regressão linear. Os resultados encontrados são mostrados na Tabela 4.2;

Tabela 4.2 – Valores das estatísticas SQR e CIB para diferentes estimativas de quebras estruturais m .

m	0	1	2	3	4	5
SQR	52457400	42141284	35876057	31895328	29955037	29144484
CIB	6976	6889	6830	6792	6780	6786

Analisando a Tabela 4.2, nota-se que o teste procura de 0 a 5 quebras estruturais na média. No entanto, como mostrado na Seção 3.3, escolhe-se o valor de m correspondente ao menor Critério de Informação Bayesiano, sendo assim $m = 4$.

O próximo passo é executar novamente o comando `breakpoints` com o número de quebras especificado. Assim, temos que as estimativas para a data das quebras são:

Tabela 4.3 – Datas das quebras identificadas pelo procedimento de Bai e Perron (1998)

Quebra	Data da quebra
Primeira	24/05/2020
Segunda	11/09/2020
Terceira	15/01/2021
Quarta	10/04/2021

Encontradas as estimativas, é calculado o intervalo de confiança para cada data. A função `confint` do software R Development Core Team (2009) aplicada ao modelo já retorna esses intervalos, sendo eles apresentados na Tabela 4.4:

Tabela 4.4 – Intervalos de confiança para às datas das quebras identificadas pelo procedimento de Bai e Perron (1998).

Data da quebra	Intervalo de Confiança
24/05/2020	22/05/2020 - 25/05/2020
11/09/2020	10/09/2020 - 12/09/2020
15/01/2021	01/01/2021 - 16/01/2021
10/04/2021	04/04/2021 - 11/04/2021

O gráfico a seguir apresenta a série do número de infectados da COVID-19 com as médias do número de casos obtidas através de cada intervalo de confiança contidos na Tabela 4.4.

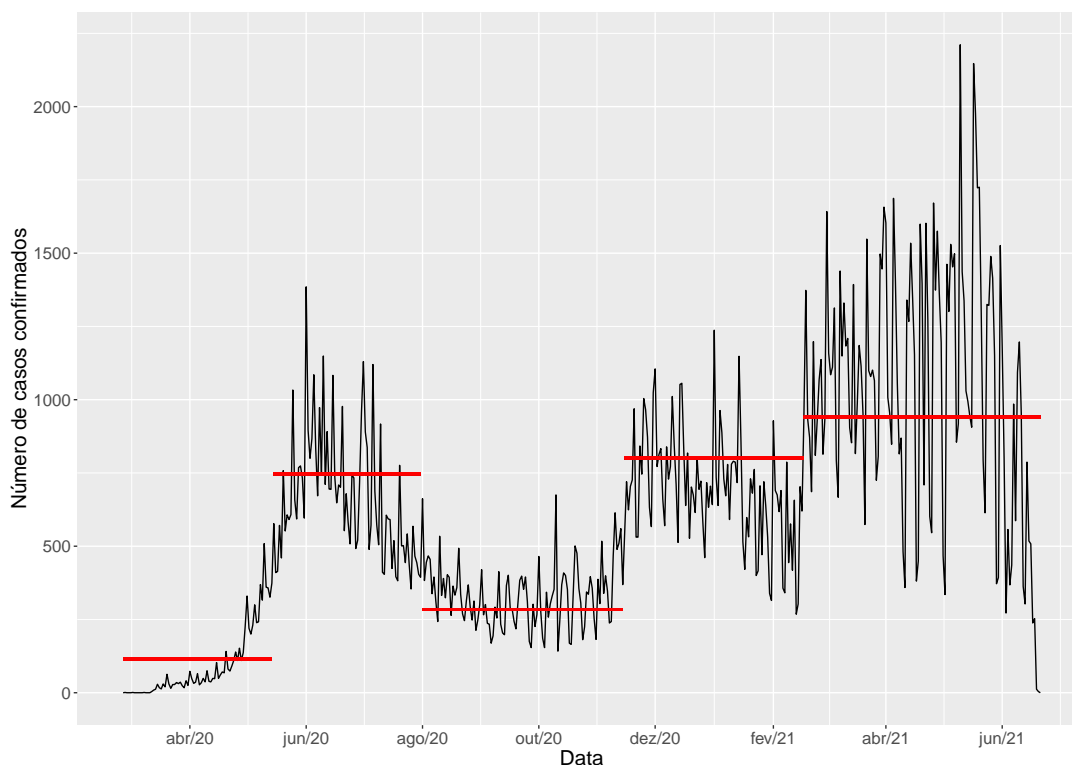


Figura 4.1 – Série do número de casos confirmados da COVID-19 no Rio Grande do Norte com estimativas das médias dos intervalos de confiança.

Os resultados apresentam que a primeira data estimada corresponde a 24/05/2020. Esta data ocorreu aproximadamente dois meses após o decreto N° 29.541, promulgado em 21/03/2020 (disponível em: Norte (2020a), acesso em: 13/07/2021). O decreto trata de medidas de lockdown. Em geral, é esperado que medidas de redução de circulação de pessoas façam com que o vírus se espalhe menos, diminuindo assim o número de infectados pela doença. Porém, como se trata do início da pandemia, é esperado que sejam detectados um grande número de casos de infectados.

Na segunda quebra, em 11/09/2020, é notada uma baixa começando no mês de julho, chegando a atingir uma estabilidade durante o início de agosto até antes do fim de novembro. Notando essa baixa em junho, no dia 19/06/2020, foi publicado o documento de N° 686500 (disponível em: Norte (2020b), acesso em: 13/07/2021) que estabeleceu a primeira fase do cronograma para a retomada gradual das atividades econômicas no Rio Grande do Norte.

No entanto, em 15/01/2021, data da terceira quebra, o número de infectados que vinha aumentando desde o final de novembro de 2020 mostra novamente uma estabilidade, mas com uma média maior do que a do período de agosto ao início de novembro, começando a aumentar novamente em fevereiro. Esse aumento pode ter ocorrido devido a aglomerações geradas no período das eleições municipais de 2020 e posteriormente no período pré-révolução como foi exibido em mídia digital (disponível em: RN (2020a), acesso em: 10/08/2021) e

(disponível em: RN (2020b), acesso em: 13/07/2021).

A quarta e última quebra foi detectada em 10/04/2021, onde o número de infectados mostra significativamente um aumento, tendo início em fevereiro e voltando a apresentar queda em meados de maio. Assim como nos anos anteriores, era previsto que o carnaval ocorresse em fevereiro, no entanto, o referido evento popular teve suas atividades proibidas em 2021, conforme o decreto estadual N° 30.369 de 1 de fevereiro de 2021 (disponível em: Norte (2021), acesso em: 03/08/2021). Porém, mesmo tendo seus eventos proibidos, em alguns lugares no Rio Grande do Norte foram constatados acontecimentos de festas que acarretaram aglomerações, como podemos constatar perante veiculações midiáticas (disponível em: RN (2021), acesso em: 13/07/2021) e (disponível em: Braziliense (2021), acesso em: 13/07/2021).

Os acontecimentos citados anteriormente podem ter sido possíveis explicações das mudanças de regime na média da série do número de casos confirmados de COVID-19 no Rio Grande do Norte, estudos como o de Shikida, Paiva e Junior (2016) obtiveram resultados de que alguns eventos históricos podem ter sido causas das oscilações na série do preço do boi gordo no estado de São Paulo no período de 1954-2012. Devido ao contágio da COVID-19 se resumir em contato de pessoas, aglomerações de tal modo podem induzir um alto grau de contágio, fazendo com que o número de infectados aumente. De modo contrário, sem aglomerações tem-se resultados diferentes, pode-se notar que um tempo depois do decreto N° 29.541 já mencionado, o número de casos de infectados apresenta uma queda, ou seja, durante o período deste decreto até o decreto N° 686500, o que nos leva a presumir que não houveram atos de aglomeração.

5 Considerações Finais

Estudos para o entendimento do surgimento e cronologia da COVID-19 estão se tornando a cada dia mais importantes, não só para compreender melhor como o vírus funciona, mas também entender a trajetória de sua propagação, como se sucedeu e o que levou a ocorrência de aumentos e decréscimos durante sua evolução.

O método proposto por Bai e Perron (1998) para detectar quebras estruturais em séries temporais identificou que as datas estimadas para as quebras existentes na série do número de infectados no Rio Grande do Norte aproximam-se a datas de eventos que possam por sua vez terem colaborado com oscilações na série. Por exemplo, na data da segunda quebra, a qual ocorreu em 11/09/2020, o estado encontrava-se em período de lockdown, por este motivo a série mostrava uma baixa seguida de uma estabilidade. No entanto, as datas estimadas para a terceira e quarta quebra ocorreram um pouco depois de alguns eventos que foram notificados como ocorrências de aglomeração, como as festividades de pré-révellion e eventos durante o período de carnaval, ou seja, possa ser que os aumentos encontrados após as quebras sejam devidos a estas ocorrências de aglomeração.

Os resultados obtidos neste trabalho mostram a necessidade da colaboração de grupos multidisciplinares para casos pandêmicos ou eventos semelhantes, onde medidas de isolamento social possam reduzir de forma considerável o alastramento de certas enfermidades. Mesmo com a escassez de trabalhos acadêmicos que servissem de referência para o desenvolvimento desta monografia, este trabalho poderá servir como base para que novos estudos sobre a COVID-19 possam surgir, gerando um melhor entendimento sobre as diferentes ondas da contaminação deste vírus.

Referências

- ABREU, C. C. A. d. Abertura econômica, eficiência técnica e produtividade total dos fatores do brasil e américa latina. 2018.
- ALVES, R.; DELGADO, C. Processos estocásticos. 1997.
- ALVES, R.; DELGADO, C. Processos estocásticos. 1997.
- BAI, J. Least squares estimation of a shift in linear processes. *Journal of Time Series Analysis*, Wiley Online Library, v. 15, n. 5, p. 453–472, 1994.
- BAI, J. Estimating multiple breaks one at a time. *Econometric theory*, Cambridge University Press, v. 13, n. 3, p. 315–352, 1997.
- BAI, J. Estimation of a change point in multiple regression models. *Review of Economics and Statistics*, MIT Press 238 Main St., Suite 500, Cambridge, MA 02142-1046, USA journals . . . , v. 79, n. 4, p. 551–563, 1997.
- BAI, J.; PERRON, P. Estimating and testing linear models with multiple structural changes. *Econometrica*, JSTOR, p. 47–78, 1998.
- BRADY, R. R. Structural breaks and consumer credit: Is consumption smoothing finally a reality? *Journal of macroeconomics*, Elsevier, v. 30, n. 3, p. 1246–1268, 2008.
- BRAZILIENSE, C. *Para governadora do RN, festas de carnaval clandestinas são 'escárnio'*. 2021. Disponível em: <<https://www.correiobraziliense.com.br/brasil/2021/02/4906830-para-governadora-do-rn-festas-de-carnaval-clandestinas-sao-escarnio.html>>.
- CAVALCANTE, J. R. et al. Covid-19 no brasil: evolução da epidemia até a semana epidemiológica 20 de 2020. *Epidemiologia e Serviços de Saúde*, SciELO Brasil, v. 29, 2020.
- CAVALCANTI, G. da S. An alise de sobrevivência dos infectados pela covid-19 no estado do rio grande do norte. 2021.
- CIOTTI, M. et al. The covid-19 pandemic. *Critical reviews in clinical laboratory sciences*, Taylor & Francis, v. 57, n. 6, p. 365–388, 2020.
- DANIEL, J. Education and the covid-19 pandemic. *Prospects*, Springer, v. 49, n. 1, p. 91–96, 2020.
- EDUCACIONAL, P. *Análise de Séries Temporais*. 2020. Disponível em: <<https://proeducacional.com/ead/curso-cga-modulo-i/capitulos/capitulo-4/aulas/analise-de-series-temporais/>>.
- EMILIANO, P. C. et al. Critérios de informação de akaike versus bayesiano: análise comparativa. *19º Simpósio Nacional de Probabilidade e Estatística*, 2010.
- ENDO, M. H.; GOMES, F. Suavização de consumo e quebras estruturais. In: ASSOCIAÇÃO NACIONAL DOS CENTROS DE PÓS-GRADUAÇÃO EM ECONOMIA. *Anais do XLIV Encontro Nacional de Economia [Proceedings of the 44th Brazilian Economics Meeting]*. [S.l.], 2018.

FARIAS, H. S. d. O avanço da covid-19 e o isolamento social como estratégia para redução da vulnerabilidade. *Espaço e Economia. Revista brasileira de geografia econômica*, Núcleo de Pesquisa Espaço & Economia, n. 17, 2020.

FIOCRUZ. *O que é uma pandemia*. 2021. Disponível em: <<https://g1.globo.com/rn/rio-grande-do-norte/eleicoes/2020/noticia/2020/09/14/convencoes-partidarias-geram-aglomeracoes-pelo-interior-do-rn.ghtml>>.

FREITAS, K. *Surto, Pandemia, Epidemia E Endemia: Compreenda A Diferença*. 2020. Disponível em: <<https://www.drakeillafreitas.com.br/surto-pandemia-epidemia-e-endemia/>>.

GONÇALVES, A. C. T. et al. Perfil epidemiológico da covid-19 no estado do rio grande do norte: análise a partir da faixa etária dos casos positivos. *Research, Society and Development*, v. 9, n. 11, p. e3539119884–e3539119884, 2020.

HALL, A. R.; HAN, S.; BOLDEA, O. Inference regarding multiple structural changes in linear models with endogenous regressors. *Journal of econometrics*, Elsevier, v. 170, n. 2, p. 281–302, 2012.

KALE, P. L.; FERNANDES, C.; NOBRE, F. F. Padrão temporal das internações e óbitos por diarreia em crianças, 1995 a 1998, rio de janeiro. *Revista de Saúde Pública*, SciELO Brasil, v. 38, p. 30–37, 2004.

LANNE, M.; LÜTKEPOHL, H.; SAIKKONEN, P. Comparison of unit root tests for time series with level shifts. *Journal of time series analysis*, Wiley Online Library, v. 23, n. 6, p. 667–685, 2002.

LANNE, M.; LÜTKEPOHL, H.; SAIKKONEN, P. Comparison of unit root tests for time series with level shifts. *Journal of time series analysis*, Wiley Online Library, v. 23, n. 6, p. 667–685, 2002.

NORTE, D. O. do Estado do Rio Grande do. *Diário Oficial do Estado do Rio Grande do Norte*. 2020. Disponível em: <http://diariooficial.rn.gov.br/dei/dorn3/docview.aspx?id_jor=00000001&data=20200321&id_doc=678003>.

NORTE, D. O. do Estado do Rio Grande do. *Diário Oficial do Estado do Rio Grande do Norte*. 2020. Disponível em: <http://diariooficial.rn.gov.br/dei/dorn3/docview.aspx?id_jor=00000001&data=20200619&id_doc=686500>.

NORTE, D. O. do Estado do Rio Grande do. *Diário Oficial do Estado do Rio Grande do Norte*. 2021. Disponível em: <http://diariooficial.rn.gov.br/dei/dorn3/docview.aspx?id_jor=00000001&data=20210202&id_doc=711847>.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2009. ISBN 3-900051-07-0. Disponível em: <<http://www.R-project.org>>.

RN, G. *Convenções Partidárias Geram Aglomerações Pelo Interior do RN*. 2020. Disponível em: <<https://g1.globo.com/rn/rio-grande-do-norte/eleicoes/2020/noticia/2020/09/14/convencoes-partidarias-geram-aglomeracoes-pelo-interior-do-rn.ghtml>>.

- RN, G. *Primeiro dia de festas pré-réveillon em Pipa e São Miguel do Gostoso tem aglomerações e falta de máscaras entre participantes*. 2020. Disponível em: <<https://g1.globo.com/rn/rio-grande-do-norte/noticia/2020/12/28/primeiro-dia-de-festas-de-reveillon-em-pipa-e-sao-miguel-do-gostoso-tem-aglomeracoes-e-falta-de-mascaras-entre-participantes.ghtml>>.
- RN, G. *PM registra aglomerações em festas clandestinas e bares são interditados no RN*. 2021. Disponível em: <<https://g1.globo.com/rn/rio-grande-do-norte/noticia/2021/02/16/pm-registra-aglomeracoes-em-festas-clandestinas-e-bares-sao-interditados-no-rn.ghtml>>.
- SHIKIDA, C.; PAIVA, G. L.; JUNIOR, A. F. A. Análise de quebras estruturais na série do preço do boi gordo no estado de são paulo. *Economia Aplicada*, Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto-USP . . . , v. 20, n. 2, p. 265, 2016.
- SIGNOR, D. *Análise de Séries Temporais*. 2020. Disponível em: <<https://proeducacional.com/ead/curso-cga-modulo-i/capitulos/capitulo-4/aulas/analise-de-series-temporais/>>.
- TIMES, T. N. Y. *Coronavirus World Map: Tracking the Global Outbreak*. 2022. Disponível em: <<https://www.nytimes.com/interactive/2021/world/covid-cases.html>>.
- WOODRIDGE, J. *Econometria: Uma abordagem moderna*. Ed. Cengage Learning. 4^a Edição, 2012.
- ZEILEIS, A. et al. Testing and dating of structural changes in practice. *Computational Statistics & Data Analysis*, Elsevier, v. 44, n. 1-2, p. 109–123, 2003.