

UNIVERSIDADE DE SÃO PAULO

ESCOLA POLITÉCNICA

RODRIGO MARCO BASSI GASPARIAN

USO DE SÉRIES TEMPORAIS E GRÁFICOS DE CONTROLE NO MONITORAMENTO
DA INFLAÇÃO

SÃO PAULO

2021

RODRIGO MARCO BASSI GASPARIAN

USO DE SÉRIES TEMPORAIS E GRÁFICOS DE CONTROLE NO MONITORAMENTO
DA INFLAÇÃO

Trabalho de formatura apresentado à
Escola Politécnica da Universidade de
São Paulo para obtenção do diploma de
Engenheiro de Produção.

Orientador: Profa. Dra. Linda Lee Ho

SÃO PAULO

2021

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catálogo-na-publicação

Gasparian, Rodrigo Marco Bassi

Uso de séries temporais e gráficos de controle no monitoramento da inflação / R. M. B. Gasparian -- São Paulo, 2021.
93 p.

Trabalho de Formatura - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Produção.

1.Séries Temporais 2.Regressão 3.Modelos para processos estacionários
4.Controle estatístico do processo 5.Inflação I.Universidade de São Paulo.
Escola Politécnica. Departamento de Engenharia de Produção II.t.

AGRADECIMENTOS

Agradeço a toda a minha família, por estarem sempre presentes em todos os momentos da minha vida, mas especialmente aos meus pais, Rubens e Adriana, por todo o esforço que colocaram na minha formação e ao meu irmão, também Rubens, por sempre me acompanhar nas mais diversas situações do dia a dia.

Agradeço a todos os meus amigos que fiz durante a graduação na Escola Politécnica e na TU Darmstadt, mas especialmente a Artur Gil, João Paulo Parizotto, Lucas Vidigal, Pedro Mello e Ricardo Brinati, que estiveram mais presentes na rotina da Poli e contribuíram imensamente para a minha formação acadêmica e profissional. Agradeço também a todos os meus amigos da vida, do ensino fundamental e médio no Colégio Visconde de Porto Seguro e de Darmstadt, que sempre me apoiaram e tornaram o dia a dia fora da universidade mais leve e divertido.

Agradeço à Profa. Dra. Linda Lee Ho por todos os ensinamentos, apoio e conselhos durante a elaboração da elaboração deste trabalho. Sem todo o acompanhamento ao longo do ano, nada disso teria sido possível. Agradeço também a todos os outros docentes que tive contato ao longo da Escola Politécnica e que contribuíram para a minha formação acadêmica.

Agradeço, por fim, aos meus colegas, mentores e gestores de todos os estágios realizei ao longo da graduação, mas especialmente à equipe da Journey Capital, que sempre me acolheu e me deu espaço e ensinamentos para me desenvolver pessoalmente e profissionalmente.

RESUMO

Ao longo da última década, o mercado norte-americano passou por um crescimento acelerado da participação de algoritmos quantitativos como estratégias de investimento, capitaneado por exemplos que obtiveram extremo sucesso através do uso de modelos estatísticos como forma de tomada de decisões. No mercado brasileiro, porém, o cenário é ainda diferente, na medida que os veículos de investimento que usam apenas estratégias quantitativas ainda não têm seu espaço consolidado no mercado e tão pouco a confiança dos investidores. Também ao longo das últimas décadas, a inflação brasileira, oficialmente o índice IPCA, mostrou-se uma preocupação recorrentes da população, ao passo que existem atualmente diversas alternativas para que o investidor consiga expor seu patrimônio em ativos indexados à índices de preços no Brasil. Diante desse contexto, esse trabalho propõe uma metodologia quantitativa para a previsão da inflação brasileira através de modelagem de séries de tempo e monitoramento por gráficos de controle. A metodologia quantitativa para a série de retornos mensais do IPCA utiliza modelos auto regressivos e de médias móveis do tipo ARIMA para ajuste da série e em seguida gráficos de controle com memória do tipo EWMA para monitorar o comportamento. Os resultados obtidos são comparados à série original e à série de expectativas do mercado para o IPCA mensal, de forma que a conclusão é satisfatória e identifica os momentos de maior estresse para a série original.

Palavras-Chave: inflação, IPCA, metodologias quantitativas, modelos estatísticos, séries temporais, ARIMA, controle estatístico de processos, gráficos de controle.

ABSTRACT

Over the last decade, the North American market has experienced a rapid growth in the participation of quantitative algorithms as investment strategies, led by examples that have been extremely successful through the use of statistical models as a form of decision-making. In the Brazilian market, however, the scenario is different, as the investment vehicles that use only quantitative strategies have neither their space consolidated in the market nor the confidence of investors. Also over the past decades, Brazilian inflation, officially the IPCA index, has shown itself to be a recurring concern for the country's population, while there are currently several alternatives for the investor to be able to invest in assets that are indexed to price indexes in Brazil. Given this context, this paper proposes a quantitative methodology for forecasting Brazilian inflation through time series modeling and control chart monitoring. The quantitative methodology for the IPCA monthly returns series uses autoregressive and moving average ARIMA models to fit the series and then control charts with memory EWMA to monitor the behavior. The results obtained are compared to the original series and to the market expectations series for the monthly IPCA, so that the conclusion is satisfactory and identifies the moments of greatest stress for the original series.

Keywords: inflation, IPCA, quantitative methodologies, statistical models, time series, ARIMA, statistical process control, control charts.

LISTA DE FIGURAS

Figura 1.1 - Patrimônio total dos fundos de investimento incentivados ao longo do tempo....	21
Figura 1.2 - Simulação da isenção fiscal em fundos de investimento incentivados.....	22
Figura 2.1 - Índice de Preços ao Consumidor Amplo (IPCA).....	30
Figura 3.1 - Evolução da participação dos <i>Quant Hedge Funds</i> no mercado dos EUA	37
Figura 3.2 - Fundos Quantitativos no Brasil.....	38
Figura 3.3 - Desempenhos dos fundos quantitativos brasileiros	39
Figura 4.1 - Fluxograma de escolha do gráfico de controle	52
Figura 5.1 - Série IPCA mensal.....	58
Figura 5.2 - Histograma da série IPCA mensal	59
Figura 5.3 - Q-Q Plot normal para a série IPCA mensal	60
Figura 5.4 - ACF da série IPCA mensal	61
Figura 5.5 - PACF da série IPCA mensal.....	62
Figura 5.6 - Séries original e modelo ARMA(2,3) com média diferente de zero	64
Figura 5.7 - Q-Q Plot normal para os resíduos do modelo ARMA(2,3) com média diferente de zero	65
Figura 5.8 - Análise ACF e PACF para os resíduos do modelo	66
Figura 5.9 – Teste Ljung-Box para os resíduos do modelo ARIMA	67
Figura 5.10 - Fluxograma para simulação e definição dos limites de controle	68
Figura 5.11 - Fluxograma da simulação do ARL	69
Figura 5.12 - Valores do ARL para variação do k	70

Figura 5.13 - Gráfico de controle para a série ARIMA	71
Figura 6.1 - Séries original e modelo ARMA(2,3) com média diferente de 0.....	73
Figura 6.2 - Simulação dos valores futuro para o modelo ARIMA da série IPCA mensal	76
Figura 6.3 – Gráfico de controle final para o modelo ARIMA da série IPCA mensal.....	77

LISTA DE TABELAS

Tabela 1.1 - CVM - Journey Capital Endurance Juros Reais	24
Tabela 3.1 - Comparação de retornos de diferentes fundos de investimento	35
Tabela 5.1 – Estatísticas descritivas de série inicial	57
Tabela 5.2 - Coeficientes do modelo ARMA(2,3) com média diferente de zero	63
Tabela 5.3 - Valores de k e ARL	70
Tabela 6.1 – Estatísticas da diferença entre as séries	74
Tabela 6.2 – Estatísticas da diferença entre as séries	75
Tabela 6.3 – Pontos fora de controle	78

LISTA DE ABREVIATURAS E SIGLAS

ACF: *Autocorrelation Function*
ADF: *Augmented Dickey-Fuller test*
AR: *Autoregressive*
ARFIMA: *Autoregressive Fractionally Integrated Moving Average*
ARIMA: *Autoregressive Integrated Moving Average*
ARL: *Average Run Length*
ARMA: *Autoregressive Moving Average*
BCB: Banco Central do Brasil
CEP: Controle Estatístico de Processos
CMN: Conselho Monetário Nacional
CNPJ: Conselho Nacional da Pessoa Jurídica
CUSUM: *Cumulative Sum*
CVM: Comissão de Valores Mobiliários
ETF: *Exchange Traded Funds*
EUA: Estados Unidos da América
EWMA: *Exponentially Weighted Moving Average*
FACV: Função de Autocovariância
FII: Fundos de Investimento Imobiliário
FIPE: Fundação Instituto de Pesquisas Econômicas
IBGE: Instituto Brasileiro de Geografia e Estatística
IGP: Índice Geral Preços
INPC: Índice Nacional de Preços ao Consumidor
IPC: Índice de Preços ao Consumidor
IPCA: Índice Nacional de Preços ao Consumidor Amplo
LC: Linha Central
LIC: Limite Inferior de Controle
LSC: Limite Superior de Controle
MA: *Moving Average*
NTN-B: Notas do Tesouro Nacional de Série B
PACF: *Partial Autocorrelation Function*

SUMÁRIO

1	INTRODUÇÃO	19
1.1	Motivação e contextualização	19
1.2	Apresentação da empresa e do objeto alvo do estudo	23
1.3	Definição do problema	24
1.4	Definição dos objetivos	25
1.5	Estrutura do trabalho	25
2	A INFLAÇÃO	27
2.1	Definição	27
2.2	Justificativa pela escolha do mercado	30
2.3	Contextualização global	32
2.4	Evolução do mercado no Brasil	32
2.5	Formas de operação no mercado	34
3	MODELOS QUANTITATIVOS NO MERCADO FINANCEIRO	35
3.1	Contextualização histórica global	35
3.2	Evolução do mercado no Brasil	37
4	REVISÃO BIBLIOGRÁFICA	41
4.1	Análise de Séries temporais	41
4.1.1	Estacionariedade, normalidade e independência	42
4.1.2	Modelos ARIMA	44
4.1.3	Resíduos	47

4.2	Controle estatísticos de processo	48
4.2.1	Gráfico de controle.....	49
4.2.2	Escolha do Gráfico de Controle	52
4.2.3	Gráfico de Controle EWMA	54
5	APLICAÇÃO PRÁTICA	57
5.1	Considerações iniciais	57
5.2	Preparativos para aplicação do modelo	58
5.3	Modelo ARIMA.....	62
5.4	Análise dos resíduos	64
5.5	Controle do modelo e geração dos sinais	67
6	ANÁLISE DOS RESULTADOS.....	73
6.1	Modelo ARIMA.....	73
6.1.1	Comparação com as expectativas do mercado.....	74
6.1.2	Simulação dos dados para períodos fora de análise	75
6.2	Gráfico de Controle	76
6.2.1	Análise do parâmetro Lambda (λ).....	77
6.2.2	Estudo dos pontos fora de controle	78
6.3	Considerações Finais.....	79
7	CONCLUSÃO	81
7.1	Desenvolvimento e resultados obtidos.....	81
7.2	Análise da adequação ao objetivo do trabalho	81

REFERENCIAS BIBLIOGRÁFICAS	83
---	-----------

APÊNDICE A: CÓDIGO EM R PARA APLICAÇÃO DO CASO PRÁTICO	87
---	-----------

1 INTRODUÇÃO

Este primeiro capítulo do Trabalho de Formatura traz o contexto em que o presente projeto está inserido, assim como dados iniciais do setor que servem de motivação para o trabalho. Além disso, são apresentados a empresa envolvida no trabalho, o objetivo alvo de estudo dentro da empresa, assim como definidos o problema e o objetivo final do projeto. Por fim, é apresentada a estrutura do trabalho de formatura.

1.1 Motivação e contextualização

Em conjunto com a Journey Capital, gestora de fundos de investimento em que o autor deste projeto é estagiário desde agosto de 2020 até o presente momento, a ser apresentada na Seção 1.2, o presente trabalho de formatura foi pensado para ajudar a enfrentar a enorme competitividade do crescente mercado de investimentos focados em ativos ligados à inflação no Brasil, principalmente os fundos de debêntures incentivadas.

Em uma breve contextualização do mercado, segundo a XP Investimentos (2021), a maior corretora de investimentos do Brasil, debêntures incentivadas são títulos de dívida emitidos por empresas não financeiras cujos recursos são destinados para obras ou serviços de infraestrutura (por exemplo, construção de rodovias, ferrovias, linhas de energia). Conforme a Casa Civil da Presidência da República (2011), por ser um assunto de interesse nacional, as debêntures incentivadas são isentas do recolhimento do Imposto de Renda, conforme a Lei Nº 12.431, de 24 de junho de 2011.

Ainda segunda a XP Investimentos (2021), as debêntures incentivadas são títulos de renda fixa – ou seja, as empresas captam recursos junto ao mercado para compor o seu passivo, e em troca pagam uma taxa de juros como recompensa pelo capital investido. Por suas características específicas, as debêntures incentivadas têm prazos longos de liquidação – os títulos são emitidos com no mínimo 4 anos para o vencimento. Além disso, são atreladas a algum índice de inflação, como o Índice Nacional de Preços ao Consumidor Amplo (IPCA) acrescido de uma taxa de juros ao ano (Cupom) e se assemelham, portanto, aos títulos do Tesouro IPCA, antigas Notas do Tesouro Nacional de Série B (NTN-Bs).

Segundo o blog da Pag Seguro (2021), outra corretora brasileira de investimentos, os fundos são como uma grande vaquinha de investidores. Eles compram cotas de participação, juntam uma quantia maior e esse dinheiro é usado para comprar diversos títulos de debêntures

incentivadas. Quem faz esse trabalho de seleção das aplicações é a equipe que administra o fundo.

Assim, os fundos de debêntures incentivadas são fundos que investem, segundo definições da Casa Civil da Presidência da República (2011), artigo 3º da Lei Nº 12.431, ao menos 85% do seu patrimônio em ativos incentivados, neste caso apenas debêntures incentivadas. Esses fundos são isentos de imposto de renda para pessoas físicas e, conforme a Pag Seguro (2021), são fundos de renda fixa, que oferecem bons rendimentos, menor risco e valor de investimento baixo, considerados assim como uma alternativa para diversificar a carteira de um investidor.

Conforme informações da XP Investimentos (2021), os fundos de debêntures incentivadas são de dois tipos: com ou sem *hedge* (entende-se por *hedge* uma proteção). Como as debêntures incentivadas possuem remuneração atrelada ao IPCA mais uma taxa pré-fixada, os fundos que possuem *hedge* têm a indexação do fundo transformada (integralmente ou parcialmente) em IPCA para uma remuneração pós-fixada. Os fundos sem *hedge*, por sua vez, possuem remuneração atrelada puramente ao IPCA.

Em números, conforme estudo realizado pela XP Investimentos (2021), com a exceção do ano de 2020 cujo volume total de emissão foi de somente R\$ 18,3 bilhões como consequência das incertezas trazidas pela pandemia - as emissões do mercado (de debêntures incentivadas) têm crescido aproximadamente 52% ao ano. No ano de 2021, ainda segundo a mesma fonte de informações, as emissões já estão aquecidas, com um total de R\$ 5,3 bilhões emitidos até fevereiro, um volume que representa mais de 2,5x os valores do mesmo período de 2020.

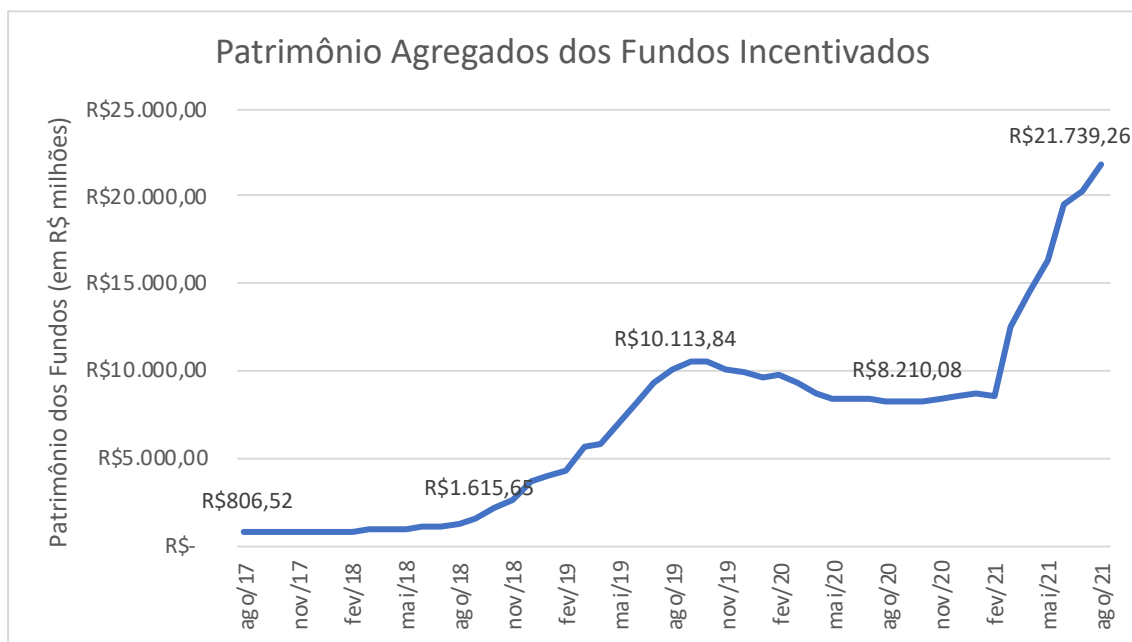
Além disso, de acordo com a XP Investimentos (2021), os fundos de investimento de debêntures incentivadas têm crescido em relação ao mercado, e nos últimos 5 anos acumulou 18,5% das emissões em média, enquanto 24,5% das emissões em média foram alocadas diretamente para pessoas físicas.

Exclusivamente em relação aos fundos de debêntures incentivadas, através de uma compilação de dados disponíveis na plataforma de análise de mercado Comdinheiro, foi possível estimar que o patrimônio dos fundos em questão subiu de R\$ 800 milhões para quase 22 bilhões nos últimos 4 anos, um aumento de mais de 25 vezes no tamanho total dos fundos.

Ainda que parte desse aumento seja proveniente da valorização das cotas dos fundos, é perceptível que há um aumento expressivo no tamanho do mercado.

O gráfico do estudo explicado está apresentado na Figura 1.1.

Figura 1.1 - Patrimônio total dos fundos de investimento incentivados ao longo do tempo



Fonte: Elaborado pelo autor (2021)

As principais justificativas para o investimento em um fundo de debêntures incentivadas, conforme artigo feito pelo blog da Pag Seguro (2021) estão listadas abaixo:

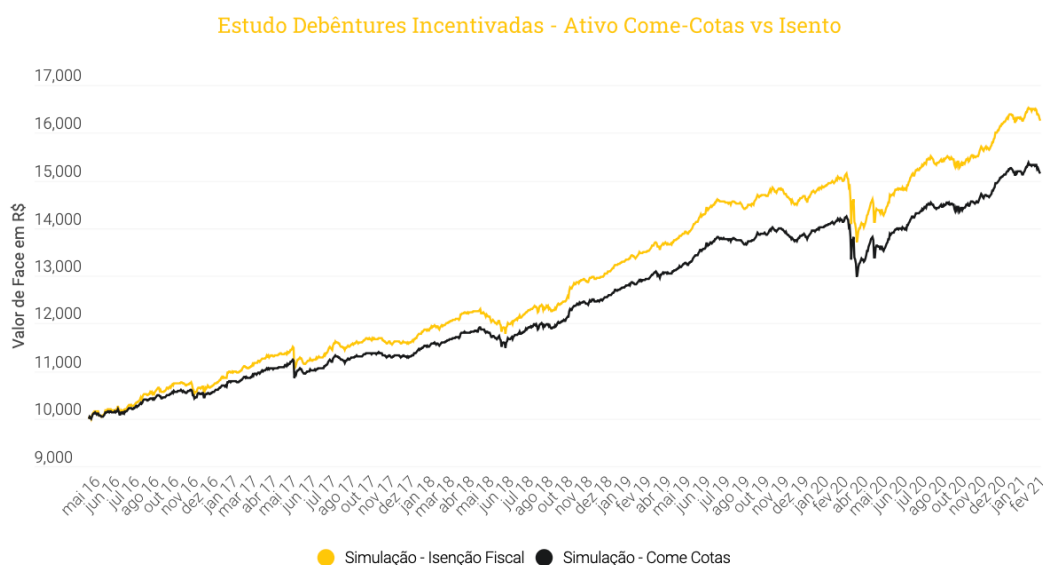
- 1) Acesso a um conjunto de debêntures incentivadas: através de uma única aplicação, o investidor acessa um conjunto de debêntures, garante-se, conforme estudo da XP investimentos, a diversificação;
- 2) Baixa taxa de administração: normalmente, a taxa de administração dos fundos de debêntures incentivadas não supera 1% ao ano, mesmo com todos os trabalhos que envolvem a gestão do fundo;
- 3) Menor aporte financeiro que o investimento direto: é possível investir em uma carteira diversificada, escolhida por um gestor profissional, por um aporte bem mais acessível. Existem aplicações mínimas de apenas R\$ 100 para estes fundos, enquanto no investimento direto o valor inicial para investir costuma ser muito mais alto (por volta de R\$ 1000);
- 4) Risco reduzido: como as debêntures são um empréstimo, o principal risco é que a empresa emissora do título não pague os juros prometidos, o chamado risco de

crédito. Ao investir em fundos de debêntures incentivadas, esse risco é pulverizado. Isso acontece porque o investidor tem acesso a um pacote de debêntures, então, se alguma empresa que compõe o fundo apresentar problemas financeiros, a perda será menor.

- 5) Gestão do fundo é profissional e otimizada: ao adquirir uma ou mais cotas de fundos de debêntures incentivadas, o investidor irá contar com uma ótima gestão de investimentos. Isso quer dizer que um gestor especialista irá gerenciar o fundo e escolher os melhores títulos do mercado para novas aplicações. O gestor faz uma avaliação bastante criteriosa para entregar ao investidor as melhores opções;
- 6) Isenção do Imposto de Renda: quem aplica dinheiro nos fundos de investimento de debêntures incentivadas recebe toda a rentabilidade, sem incidência de imposto sobre o rendimento, como acontece em outros fundos, como por exemplo os multimercados.

A partir da última vantagem mostrada, a XP Investimentos (2021) realizou um estudo que tentou mostrar na prática quais seriam os ganhos tributários dessa isenção de IR que é vantagem dos fundos de debêntures incentivadas. Em uma simulação, a XP Investimentos (2021) levou em conta os seguintes parâmetros: ganho tributário ao longo de dez anos que um fundo isento tem em relação a um fundo de renda fixa tradicional. O resultado está mostrado na Figura 1.2.

Figura 1.2 - Simulação da isenção fiscal em fundos de investimento incentivados



Fonte: XP Investimentos (2021)

Percebe-se pela Figura 1.2, que a curva de cor preta no gráfico, referente a simulação Come-Cotas, ou seja, não isenta de tributação, teve um valor de face que iniciou em R\$ 10.000 e finalizou o período de 5 anos em aproximadamente R\$ 15.000, o que indica uma rentabilidade próxima de 50%. Enquanto isso, a linha amarela do gráfico, representada pela simulação da Isenção Fiscal, portanto isenta de tributação, saltou de R\$ 10.000 no valor de face para aproximadamente R\$ 16.000, um aumento de 60%, um retorno cerca de 10% maior que a simulação tributada.

Dada a motivação e contextualização apresentadas acima, o projeto tem como inspiração apresentar um modelo quantitativo com uma base teórica robusta, resultados consistentes e um plano de ação para controle do modelo para que possa ser usado na prática pela gestora de fundos de investimento.

1.2 Apresentação da empresa e do objeto alvo do estudo

A Journey Capital Administração de Recursos Ltda., registrada sob Conselho Nacional da Pessoa Jurídica (CNPJ) número 20.316.689/0001-75 é uma sociedade que tem como atividade principal a gestão de fundos de investimento. A gestora existe no formato atual desde meados de 2017 e é composto por cinco principais sócios, todos com vasta experiência no mercado de capitais ao longo dos últimos trinta anos até se juntarem para formar a Journey Capital no seu atual formato.

O principal produto financeiro trabalhado pela gestora são créditos privados, que são títulos de renda fixa emitidos por instituições financeiras ou empresas, mas principalmente as debêntures incentivadas, uma forma específica do produto definido. Segundo a Journey Capital (2019), é nesse ramo do mercado de renda fixa que são investidos a maior parte do capital humano e intelectual da sociedade, o que leva ao mercado cinco diferentes fundos de investimento, sendo um deles objeto alvo de estudo deste presente trabalho de formatura: o *Journey Capital Endurance Juros Reais Master Fundo Incentivado de Investimento em Infra Renda Fixa* registrado sob CNPJ número 34.706.691/0001-65.

Conforme a Comissão de Valores Mobiliários (CVM) (2019), o Endurance Juros Reais é um fundo de investimento incentivado, portanto, enquadrando-se nas definições expostas na Seção 1.1, ao alocar ao menos 85% do seu patrimônio em debêntures incentivadas indexadas à inflação. O fundo de investimento está registrado desde setembro de

2019 e está atualmente em funcionamento normal, conforme informações da Tabela 1.1, retirada da CVM (2019).

Tabela 1.1 - CVM - Journey Capital Endurance Juros Reais

Consulta Consolidada de Fundo	
Nome do Fundo:	CNPJ:
JOURNEY CAPITAL ENDURANCE JUROS REAIS MASTER FUNDO INCENTIVADO DE INVESTIMENTO	34.706.691/0001-65
Administrador:	CNPJ:
BTG PACTUAL SERVIÇOS FINANCEIROS S/A DTVM	59.281.253/0001-23
Situação Atual:	Data Início de Atividades
EM FUNCIONAMENTO NORMAL	18/09/2019
Data de Constituição	Web Site:
05/09/2019	WWW.BTGPACTUAL.COM

Fonte: Elaborado pelo autor (2021)

Conforme a Journey Capital (2019), a Journey Capital Administração de Recursos é apresentada como gestor do fundo de investimento, o que dá poderes para negociar, em nome do fundo, os ativos financeiros que compõem a sua carteira.

Por fim, conforme a CVM (2019), no último dia útil de agosto de 2021, o fundo apresentava um patrimônio líquido de R\$ 22,626,194.27 e uma cota de 1.235481.

1.3 Definição do problema

No contexto diário de gestão do Endurance Juros Reais, que busca entregar aos seus investidores uma rentabilidade maior que o seu *benchmark* e acima de seus concorrentes, é preciso ter métricas e modelos bem definidos para que a alocação dos ativos seja feita da melhor forma possível conforme as diferentes situações do mercado.

Porém, conforme destacado anteriormente, em um mercado que tem um crescimento acelerado como este, a competição é intensa e repleta de *players* conhecidos e consolidados no mercado. Enquanto existem gestoras independentes do mercado que empregam mais de 100 pessoas dedicadas à análise dos ativos e tomada de decisões de seus fundos, a Journey conta com uma equipe de apenas 15 pessoas.

Em outras palavras, há uma dificuldade enorme em trazer rentabilidade para um fundo de investimento quando concorrentes de mercado contam com uma estrutura muito mais robusta em todos os aspectos tradicionais possíveis. É por esse motivo que a gestora busca implementar no dia a dia do fundo estratégias que possam diferenciar-se do mercado, de

forma a trazer formas de raciocínio que utilizem algoritmos estatísticos, por exemplo, para além de outros métodos já utilizados.

1.4 Definição dos objetivos

Os objetivos são: propor um modelo de previsão para a série dos retornos mensais do IPCA, que é o índice oficial de inflação do Brasil, através de modelos de série temporais, como os modelos auto regressivos integrados de médias móveis (ARIMA), e monitorá-la através de um gráfico de controle, tal que seja capaz de indicar os momentos em que o preço do ativo controlado indique sinais de instabilidade do mercado, o que indica que a inflação esperada pode ter se alterado.

Ao final, deseja-se obter um modelo que gere sinais confiáveis para um melhor controle do IPCA e que possa ser implantado no dia a dia da gestora como forma de auxiliar na tomada de decisões ligadas à inflação, índice que é indexador de ao menos 85% dos ativos da carteira do Endurance Juros Reais.

1.5 Estrutura do trabalho

O presente trabalho de formatura irá estruturar-se com uma apresentação no Capítulo 2 sobre o mercado que envolve o índice a ser estudado, através de uma análise do seu histórico, importância e comparações. Além disso, não só as diferentes formas de se operar nesse mercado serão brevemente apresentadas, como também os títulos mais e menos populares existentes.

No Capítulo 3, será feita uma análise quantitativa e qualitativa de modelos estatísticos quantitativos dentro do mercado financeiro. Por tratar-se de uma estratégia de investimento relativamente nova, há um enorme potencial de crescimento, boa parte potencializada pelo já sucesso feitos nos mais diferentes mercados.

No Capítulo 4, será resumida a revisão bibliográfica que embasará o projeto, assim como serão definidas e apresentadas as metodologias de análise de séries temporais e de gráficos de controle, presente na bibliografia resumida, que irão permear a aplicação do caso prático.

No Capítulo 5 será desenvolvido um modelo prático de aplicação de todos os conceitos de séries temporais apresentados, monitorando os resultados através de gráficos de controle. Em seguida, o Capítulo 6 trará uma análise dos resultados obtidos no caso prático.

Por fim, o Capítulo 7 concluirá o trabalho com um resumo do presente trabalho de formatura e verificando se os objetivos estabelecidos foram atingidos pelos resultados obtidos.

2 A INFLAÇÃO

Este capítulo visa trazer uma visão sobre o conceito de inflação, índice que serve como base de estudo do presente trabalho de formatura. Além de uma definição detalhada do índice conforme instituições regulatórias do Brasil, serão apresentados estudos que mostram a relevância do estudo do tema no mundo e no Brasil. Por fim, será dada uma breve introdução das possíveis formas de investimento na prática em mercados ligados à inflação.

2.1 Definição

A inflação, objeto de estudo do presente estudo, é, segundo o Banco Central do Brasil (BCB) (2021), o aumento dos preços de bens e serviços e implica diminuição do poder de compra da moeda. Em outras palavras, a inflação é uma medida feita por centros de pesquisa que buscam mostrar a média de aumento do preço de uma determinada cesta de produtos e serviços. Com o passar do tempo, a inflação mostra que uma quantidade X de dinheiro irá ter um poder de comprar menor com o passar do tempo, à medida que o preço médio dos produtos e serviços da sociedade são elevados.

Ainda conforme o BCB (2021), a inflação pode ter diversas causas, que basicamente podem ser agrupadas em: pressões de demanda, quando há demanda excessiva por algum produto ou serviço, o que faz com que o seu preço aumente; pressões de custos, quando há uma oferta escassa de algum produto ou serviço, o que faz com o seu preço aumente; inércia inflacionária, basicamente a tendência natural de aumento dos preços; e expectativas de inflação, quando perspectivas futuras de um aumento de preço fora do normal, o que faz com que isso se reflita nos preços atuais.

Além disso, a inflação, segundo o BCB (2021), quando acima daquilo que é esperado pela economia, traz consigo consequências como:

- 1) Incertezas na economia, desestímulo do investimento e consequente prejuízos para o crescimento econômico;
- 2) Distorção dos preços, que geram ineficiências na economia e perda de noção dos preços relativos;
- 3) Aumento do custo da dívida pública, que é em parte remunerada através da inflação.

A inflação é medida pelos índices de preços, que, conforme o BCB (2021), nada mais são do que a média de diversos preços de modo a resumi-los em um único número. Conforme o BCB (2021), esses índices podem diferir principalmente na cesta de bens e serviços que é tomada como referência para criar a média e no período em que são coletados os preços. Conforme exemplo dado pelo BCB (2021), dois índices de preço que são voltados para os consumidores podem ainda assim diferir entre si, pois cada família tem sua própria cesta de consumo, que refluem um custo de vida diferente para cada grupo adotado como referência.

Conforme o BCB (2021), no Brasil existem quatro principais índices de preços calculados por três centros de pesquisa, cada um com suas particularidade e importância, que são:

- 1) IPCA: O Índice Nacional de Preços ao Consumidor Amplo é calculado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), mede o preço de uma cesta de consumo representativa para famílias com renda de um a quarenta salários-mínimos, em treze áreas geográficas de maior relevância no Brasil;
- 2) Índice Geral de Preços (IGP): feito pela Fundação Getúlio Vargas e calculado de uma forma mais abrangente, inclui não só preços ao consumidor, mas também preços ao produtor e custos da construção. O índice possui três versões diferentes em função do período de coleta: IGP-DI, IGP-M e IGP-10;
- 3) Índice Nacional de Preços ao Consumidor (INPC): tem cálculo similar ao IPCA, porém apenas abrange famílias com renda de um a cinco salários-mínimos. Também é calculado pelo IBGE;
- 4) Índice de Preços ao Consumidor (IPC): calculada pela Fundação Instituto de Pesquisas Econômicas (FIPE) é um dos mais antigos índices de preço ao consumidor brasileiro. A abrangência é de famílias com renda de um a dez salários-mínimos somente no município de São Paulo;

A partir de um índice de preço, a inflação pode ser definida como a variação do custo da cesta de produtos que envolve o cálculo do índice durante um período determinado.

Conforme o BCB (2021), o IPCA é o índice de referência utilizado no sistema de metas para a inflação do Brasil. Sendo assim, o Banco Central do Brasil trabalha para que a inflação anual, medida através do IPCA, se situe em torno do centro da meta de inflação definida pelo Conselho Monetário Nacional (CMN).

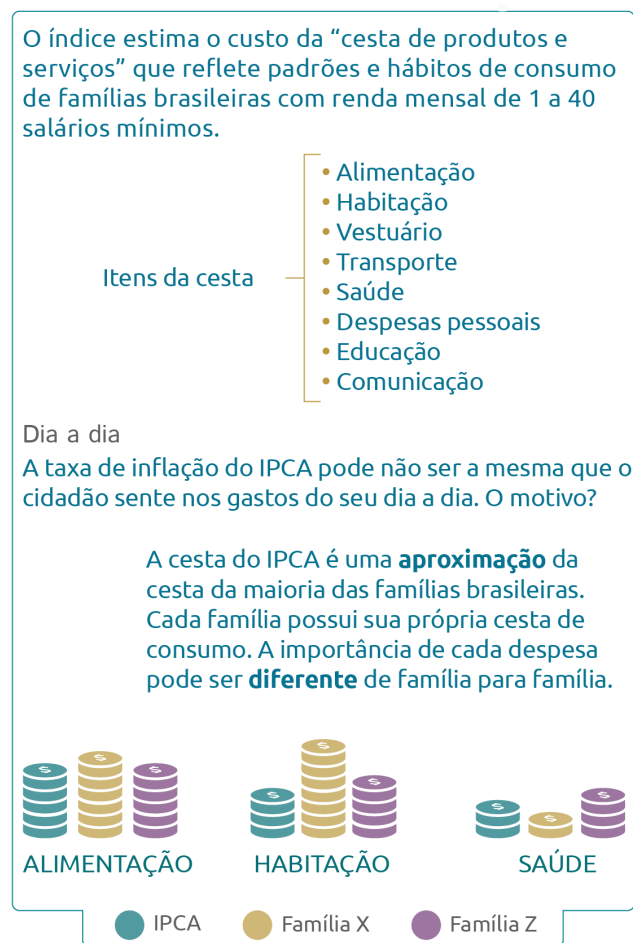
Conforme artigo do BCB (2021), a inflação oficial do Brasil é calculada pela variação do custo da cesta do IPCA durante um período determinado. Alguns itens inclusos no índice do IPCA são: alimentação, habitação, vestuário, transporte, saúde, despesas pessoais, educação e comunicação, cada um com seu peso específico dentro da média feita para a obtenção do número final.

Conforme explicado anteriormente, a inflação calculada através da variação do IPCA pode não ser percebida no dia a dia de todos os cidadãos. Isso acontece, pois a cesta de produtos é uma aproximação daquilo que é consumido pelas famílias brasileiras, pode-se então encontrar famílias que dão um peso maior para um item que receba um peso menor no cálculo do índice de preços ao consumidor amplo.

Um fato curioso, conforme o BCB (2021), é que parte da credibilidade recebida pelo IPCA vem do fato do índice não ser apurado diretamente pelo Banco Central. Sendo assim, o IBGE realiza os procedimentos de cálculo que conferem uma maior confiança para que o BCB trabalhe sempre para orientar suas ações de controle da inflação.

Figura 2.1 ilustra algumas das informações explicadas sobre o IPCA.

Figura 2.1 - Índice de Preços ao Consumidor Amplo (IPCA)



Fonte: Banco Central do Brasil (2021)

2.2 Justificativa pela escolha do mercado

A escolha pelo estudo da inflação no presente trabalho de formatura perpassa por três principais motivos, a serem explicados individualmente durante a presente Seção do trabalho de formatura.

Conforme apresentado anteriormente, a gestora em que o autor do presente projeto é estagiário é focada em fundos de debêntures incentivadas, ativos que são indexados ao IPCA. Dados os fatos em consideração, é evidente que as flutuações do IPCA mês a mês impactam diretamente as posições do fundo de investimento que é objeto alvo do presente projeto.

Portanto, a primeira justificativa pela escolha da inflação como o índice alvo de mercado é estudar um objeto que tenha alguma possibilidade real de ajuda na prática para a gestora. Já existem dentro da gestora outras formas de estudo e previsão do IPCA, porém,

todos dependem de processos manuais e subjetivos, que se baseiam muito em opiniões sobre cenários macroeconômicos.

A ideia de propor um modelo quantitativo, que consiga colaborar de alguma forma para os estudos já existentes na gestora sobre o índice é muito bem recebida pelos membros da Journey Capital.

Outro ponto positivo sobre a escolha da inflação como objeto de estudo é que, por ser um mercado em que já existe familiaridade por parte do autor, há uma facilidade maior em saber se os resultados a serem obtidos e analisados estão coerentes e se fazem algum sentido prático.

Caso fosse escolhido um objeto de mercado em que não há conhecimento prévio algum, existiria uma possibilidade maior dos resultados do projeto não serem interpretados de forma correta.

Por último, ao mesmo tempo em que é um índice antigo, consolidado e famoso no mercado brasileiro, também é um tema muito complexo, alvo de estudo de diversas pesquisas macroeconômicas, inclusive de estudo em um dos modelos mais importante da macroeconomia mundial, a curva de Phillips.

Conforme Sachsida (2013), atual Secretário de Política Econômica da Secretaria Especial da Fazenda do Ministério da Economia do Governo Brasileiro, mesmo a inflação sendo um índice ensinado amplamente em qualquer curso de economia, as pesquisas de Sachsida (2013) colocam dúvidas sobre a adequação da curva de Phillips para descrever a dinâmica inflacionária da economia brasileira.

Ainda conforme Sachsida (2013), ainda que muitos autores busquem estimar a curva de Phillips para a economia brasileira, as conclusões nunca são iguais. As diferenças são dadas principalmente devido à sensibilidade quanto ao período analisado, aos conjuntos de aproximações utilizadas e ao método de cálculo escolhido. Este autor ainda conclui, que é possível ser o momento para que outras formas de descrever a dinâmica de inflação brasileira sejam descritas.

Portanto, em um contexto que o estudo da inflação encaixa-se como um tema atual, importante, com espaço para novos estudos, com grande relevância para a gestora da qual o

autor do projeto faz parte e que, ainda assim, é familiar ao autor, a escolha do estudo do mercado inflacionário é adequada para o presente projeto.

2.3 Contextualização global

Em uma breve contextualização global do mercado inflacionário, Holland e Mori (2010), buscam discutir como o estudo da inflação mudou no período recente e como que as dinâmicas globais afetam a dinâmica da inflação também no Brasil.

Conforme estes autores, o final do século XX foi um período em que os países estiveram acostumados com taxas de inflação baixas e estáveis aliadas a um bom crescimento das economias. Tal padrão afetou também a dinâmica inflacionária brasileira, muito por conta do aprofundamento da globalização das atividades produtivas, das expectativas inflacionárias baixas, entrou outros.

Ainda conforme Holland e Mori (2010), sobre o fenômeno apresentado, existem diversas tentativas de descrição da curva de Phillips para qualquer país, mas todas esbarram nos cálculos e não se tem uma resposta adequada. De qualquer forma, o período serviu como forma de aumentar a credibilidade de diversos Bancos Centrais ao redor do mundo.

Esse aumento recente na credibilidade dos Bancos, refletiu, por exemplo, na independência que diversos bancos centrais têm hoje em dia, trabalhando sobre um regime de metas para a explicação definidos por ele mesmo com a ajuda de diversos modelos e premissas assumidas.

2.4 Evolução do mercado no Brasil

A dinâmica inflacionária na economia brasileira nem sempre foi controlada e estável como está atualmente, conforme contam Pastore e Pinotti (1998) em seu estudo sobre a inflação e a estabilização que acontece no final do século XX no Brasil, mais especificamente a partir de 1994.

Conforme os autores, na época o país encontrava-se em uma inércia inflacionária, em que havia choques constantes nas curvas de oferta e demanda, que aliada à passividade monetária e à indexação generalizada dos preços, gerava efeitos inflacionários que não se dissipavam. Ainda adiante, pode-se afirmar que, naquele momento, era necessário atacar dois pontos diferentes para conter a inflação: primeiro, uma reforma monetária que alterasse o

regime monetário e cambial; em segundo, uma reforma fiscal, de forma a conter os gastos públicos para regras orçamentárias para o governo.

Nesse contexto que, após quase 5 anos de um cenário descontrolado, foi implementado o Plano Real, feito em junho de 1994. Conforme Pastore e Pinotti (1998), o plano funcionou, gerou condição para tornar a oferta agregada vertical e trouxe a inflação para níveis mais baixos.

Por outro lado, o Plano não contemplou a reforma fiscal, que, segundo Pastore e Pinotti (1998), era o segundo ponto de maior importância. Em vez disso, criou-se uma política mais expansionista, que elevou consequentemente o déficit das contas públicas, e manteve aceso ainda o problema com a inflação.

A partir de 1999, por outro lado, iniciou-se o que foi chamado de nova política monetária, conforme Arestis, de Paula e Ferrari-Filho (2009). O regime de metas de inflação no Brasil, como é conhecido e existente até hoje, enfrentou em seu início problemas, mas que com o tempo levou a uma estabilização e melhor no nível do cenário inflacionário do país.

Porém, ainda assim, conforme Arestis, de Paulo e Ferrari-Filho (2009), ainda não se sabe ao certo qual a eficiência do plano de metas para a inflação, dado que países que enfrentaram experiências similares de preços altos conseguiram por estabilizar o nível de preço mesmo sem a implementação de um plano de tal tipo.

Ainda adiante, Arestis, de Paula e Ferrari-Filho (2009) propõem que a nova política monetária definida ainda gera preocupações. Isso acontece, pois o resultado esperado da inflação está acima do planejado enquanto o crescimento um pouco abaixo, cenário que inclusive pode ser visto nos dias de hoje de São Paulo.

Dado o contexto de evolução do mercado e o panorama atual, é natural que a busca por títulos indexados ao IPCA seja alta, principalmente no histórico recente. Nesse sentido Valor Investe (2021) relatou que a disparada da inflação no último ano levou as pessoas a buscar uma proteção contra ela no mercado de investimentos.

É natural que, para quem acompanhou toda a evolução e instabilidade da dinâmica inflacionária brasileira, exista um temor maior quando choques (mesmos que temporários como acredita-se que seja o atual) aparecem na realidade brasileira.

2.5 Formas de operação no mercado

Nesta contextualização sobre a inflação, pouco se falou, porém, como de fato são levados na prática os investimentos que buscam entregar rentabilidade superiores à da inflação, protege os investidores do possível descontrole da mesma.

Além das debêntures incentivadas, títulos de crédito privado indexados a inflação e já apresentados anteriormente, a Info Money (2021) traz algumas opções de investimento atrelados à inflação para proteger o investidor, como:

- 1) Títulos Públicos IPCA+: antigamente conhecidos como NTN-B e atualmente conhecido como Tesouro IPCA+, o papel que combina uma parte de um retorno pré-fixado (ligado a parte do “+”) e outra indexada à inflação (ligada a parte do “IPCA”);
- 2) Fundos de Inflação: existem fundos de investimento focados em ativos de renda fixa que alocam diferentes títulos indexados em sua carteira, de forma a oferecer ao investidor um bom portfólio para defender-se da inflação;
- 3) *Exchange Traded Funds* (ETF) de Renda Fixa: Fundos Negociados na Bolsa, em tradução livre, são ativos de renda fixa que buscam replicar o desempenho de índices de mercado, geralmente algum relacionado à inflação;
- 4) Fundos de Investimento Imobiliários (FII): conhecidos como FII, também pode ter em sua carteira somente ativos indexados à inflação, oferece uma proteção para o investidor.
- 5) Debêntures: não somente as incentivadas, mas também outras debêntures indexadas à índices de preços inflacionários, como IGP-M. As debêntures funcionam da mesma forma que o Tesouro IPCA+, com a uma parcela de sua remuneração pré-fixada, geralmente um pouco maior que a parcela fixada do título público, e a outra parcela indexada ao IPCA, mas pode ser também a outros índices de preços.

3 MODELOS QUANTITATIVOS NO MERCADO FINANCEIRO

O presente capítulo tem o objetivo de mostrar um breve histórico do recente ramo do mercado financeiro que envolve a aplicação de modelos quantitativos como estratégias de investimento, motivação para que o presente projeto fosse desenvolvido.

3.1 Contextualização histórica global

Zuckerman (2019) relatou que a popularização dos modelos quantitativos dentro do mercado financeiro aconteceu há pouco mais de quarenta anos, quando Jim Simons, um matemático, tomou a decisão de largar a carreira acadêmica e transitar para o mercado de investimentos e ficou conhecido depois como um dos mais famosos *traders* da história recente do mercado financeiro.

Em *The Man Who Solved the Market: How Jim Simons Launched the Quant Revolution*, Zuckerman (2009) escreve a história detalhada de como a *Renaissance Technologies Corporation*, empresa fundada por Jim Simons, dominou o mercado de investimentos após o lançamento de seu fundo de investimentos modelo *Medallion*. A estratégia do fundo, liderada por Simons a partir de 1990, era puramente baseada em modelos matemáticos e uso intensivo de dados.

Através de uma compilação de dados históricos da época, Zuckerman (2009) mostrou como a estratégia mostrada por Simons revolucionou o mercado e entregou resultados superiores a todos as mais famosas figuras do mercado financeiro, como *George Soros*, *Steve Cohen*, *Peter Lynch*, *Warren Buffet* e *Ray Dalio*.

Tabela 3.1 mostra uma comparação feita por Zuckermann (2009) que mostra a rentabilidade anualizada dos principais veículos de investimento geridos por Jim Simons, que se baseava em modelos matemáticos, e por outros personagens históricos da era recente do mercado.

Tabela 3.1 - Comparação de retornos de diferentes fundos de investimento

Comparações de Retorno			
Investidor	Fundo/Veículo de Investimento	Período	Rentabilidade anualizada
Jim Simons	<i>Fundo Madallion</i>	1988-2018	39,1 %

George Soros	<i>Fundo Quantum</i>	1969-2000	32 %
Steve Cohen	<i>SAC</i>	1992-2003	30 %
Peter Lynch	<i>Fundo Magellan</i>	1977-1990	29 %
Warren Buffett	<i>Berkshire Hathaway</i>	1965-2018	20,5 %
Ray Dalio	<i>Pure Alpha</i>	1991-2018	12 %

Fonte: Elaborado pelo autor (2021)

Após a revolução descrita no livro, a ascensão dos modelos quantitativos fez com que a estratégia ganhasse um espaço significativo no mercado financeiro dos Estados Unidos da América (EUA), o que representou, conforme *The Wall Street Journal* (2018), quase um terço do total das operações no mercado de ações da época.

Os *Quant Hedge Funds*, como também é classificado o veículo de Jim Simons, são aqueles que dependem de modelos computacionais ao invés de métodos de pesquisas tradicionais e/ou intuições. Figura 3.1, retirada do *The Wall Street Journal* (2018), mostra a evolução do mercado, que dobrou de tamanho nos últimos 8 anos e ultrapassou o volume em negociações do mercado de varejo.

Figura 3.1 - Evolução da participação dos *Quant Hedge Funds* no mercado dos EUA

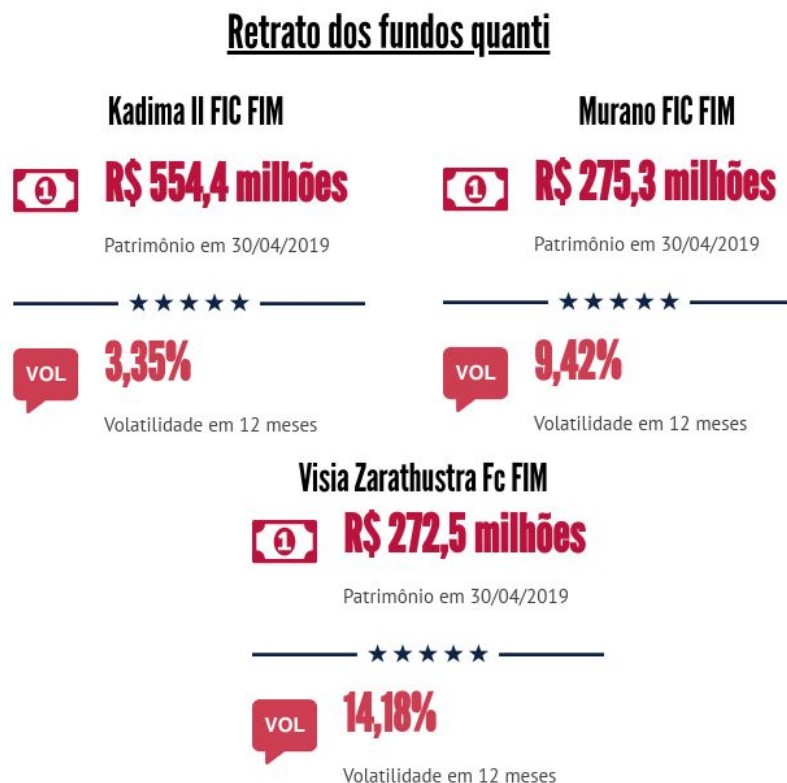


Fonte: The Wall Street Journal (2018)

3.2 Evolução do mercado no Brasil

Apesar da evolução e da popularidade dos modelos computacionais e dos fundos que usam tal estratégia de investimento nos Estados Unidos, a realidade no Brasil ainda é diferente. Conforme a *InfoMoney* (2021), há apenas três gestoras que são reconhecidamente detentoras de fundos quantitativos, e juntas somam em seus principais veículos de investimento pouco mais de um bilhão de reais, valor inexpressivo perto do volume financeiro total presente nos fundos multimercado, classificação em que são encaixados tais fundos no Brasil. Figura 3.2, retirada da reportagem, mostra o cenário que se encontram cada um dos três principais fundos quantitativos acima citado.

Figura 3.2 - Fundos Quantitativos no Brasil

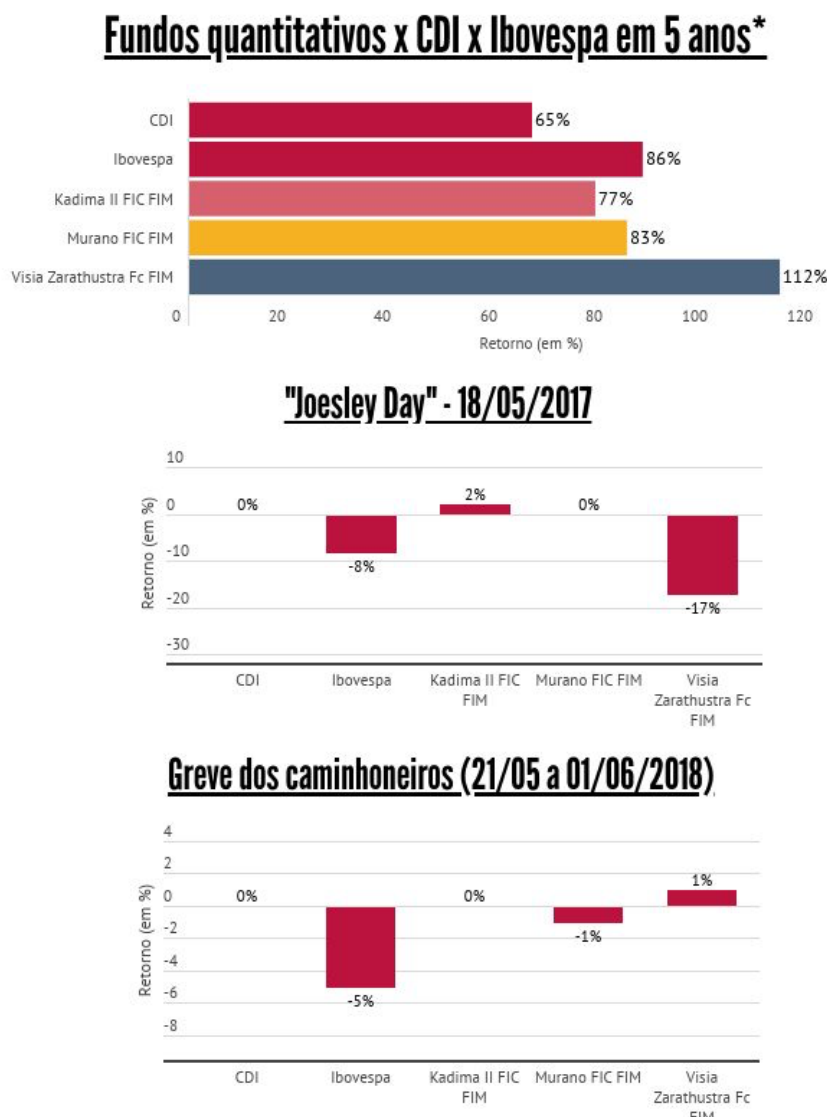


Fonte: InfoMoney (2019)

Ainda conforme o artigo, as gestoras tiveram o início de sua operação há pouco mais de dez anos e acumulam desde então resultados expressivos contra os principais *benchmarks* do mercado nacional. Uma característica marcante para tais fundos no Brasil é a proposta de se firmar como uma alternativa para diversificação das carteiras, preparados para enfrentar os principais momentos de volatilidade do mercado.

Figura 3.3 resume algum dos principais números dos três fundos destaque da reportagem. É interessante notar que o desempenho marcante do dia dezoito de maio de 2017 para o fundo *Zarathustra*, também mostrada abaixo, fez com que os modelos tirassem de aprendizado a situação, de forma que foram definidos limites diferentes para a operação, que limitam máximas perdas futuras, por exemplo.

Figura 3.3 - Desempenhos dos fundos quantitativos brasileiros



Fonte: InfoMoney (2019)

Na reportagem é citado também que para um perfil moderado de investimento no Brasil, atualmente é comum que sejam alocados em fundos quantitativos apenas de dois a cinco por cento do patrimônio total investido.

Ceciliano (2017) tentou procurar razões para a falta de procura dos fundos de investimento quantitativo no Brasil, ao analisar alguns veículos selecionados do mercado e comparar os resultados com *benchmarks* tradicionais do mercado.

A conclusão foi de que, apesar da maioria dos fundos superarem os parâmetros de mercado nos diferentes períodos analisados, o resultado mostrado ainda assim não foi bom, na medida que apresenta um desempenho muito pouco acima das expectativas do mercado.

Nota-se, portanto, que ao contrário do cenário encontrado nos Estados Unidos, relatado por Zuckermann (2019), o Brasil ainda enfrenta desafios no ramo dos modelos quantitativos, de forma que as estratégias quantitativas têm um papel pouco expressivo perto de todo o potencial que já acontece fora do país.

Por um lado, esse pequeno espaço ocupado por tais fundos no mercado abre espaço para que exista um crescimento acelerado em caso de criação de um modelo ideal que entregue resultados consistentes. Porém, por outro lado, mostra-se a dificuldade de desenvolvimento de modelos de sucesso, dado que as primeiras gestoras existem há mais de quarentena anos nos Estados Unidos e mais de dez anos no Brasil e até então nenhum resultado de expressivo destaque foi entregue no mercado brasileiro de investimentos.

4 REVISÃO BIBLIOGRÁFICA

Este capítulo visa trazer uma breve revisão das principais bibliografias que envolvem os dois principais conceitos que envolvem a análise prática a ser feita ao final do Trabalho de Formatura: análise de séries temporais e controle estatístico de processos, com foco maior em gráficos de controle.

4.1 Análise de Séries temporais

Uma série temporal foi definida por Morettin e Toloi (2006) através de “qualquer conjunto de observações ordenadas no tempo”. Como exemplo de séries temporais, tem-se entre diversos outros, os citados abaixo:

- 1) valores diários de poluição na cidade do Rio de Janeiro;
- 2) valores mensais de temperatura na cidade de Santo Antônio do Pinhal-SP;
- 3) índices diários da Bolsa de Valores de Nova Iorque-EUA;
- 4) índice mensais de atividade econômica do Brasil;
- 5) registro de marés no porto de Vitória;

É importante notar ainda que, conforme observado, as séries temporais não se limitam a nenhum campo de estudos, de forma que há aplicações tanto dentro do mercado financeiro, caso do presente Trabalho de Formatura, como em outros campos de estudos como ciências biológicas, meteorologia, ciências exatas, entre outros.

Ainda conforme Morettin e Toloi (2006), séries temporais podem ser discretas, como no caso dos exemplos 1) a 4) acima, ou contínuas, como no caso do exemplo 5) logo acima.

Para o caso de séries temporais contínuas, é comum que se faça uma conversão através de amostragens em intervalos de tempo iguais e pré-definidos para que o estudo delas seja possível. Assim, para o último exemplo acima, uma possível forma de amostragem seria medir intervalos de tempo Δt e assim obter uma série temporal discreta derivada da série contínua. Matematicamente, a equação (4.1) mostra como a conversão seria feita na série contínua de intervalo fechado $[0, T]$ e N pontos amostrados:

$$N = \frac{T}{\Delta t} \quad (4.1)$$

Mais adiante, Morettin e Tolo (2006) explicam os dois enfoques usados na análise de séries temporais, ambos com o mesmo objetivo de “construir modelos para as séries com propósitos determinados”:

- (i) domínio temporal: modelos propostos são modelos paramétricos, com um número finito de parâmetro definidos – os modelos ARIMA, a serem explicados posteriormente, e usados na parte prática do presente trabalho, são exemplos deste tipo;
- (ii) domínio de frequências: modelos propostos são modelos não- paramétricos.

No enfoque de estudo de análise de séries temporais do presente Trabalho de Formatura, além do estudo dos modelos ARIMA, é importante destacar suposições e validações feitas previamente ao modelo, assim como as análises posteriores em cima dos resíduos do modelo. Todos esses pontos serão abordados nas Seções 4.1.1, 4.1.2 e 4.1.3.

4.1.1 Estacionariedade, normalidade e independência

No estudo de séries temporais, conforme Morettin e Tolo (2006), é necessário introduzir três suposições simplificadoras:

- 1) processos estacionários ou não estacionários, conforme independência em relação à origem dos tempos;
- 2) processos normais ou não-normais, conforme as funções de densidade de probabilidade que caracterizam o processo;
- 3) processos Markovianos ou não-Markovianos, conforme independência dos valores do processo de seus valores precedentes;

De acordo com Morettin e Tolo (2006), a estacionariedade de uma série temporal é assumir que “ela se desenvolve no tempo aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável”. Intuitivamente, isso significa que um processo é estacionário se, ao se desenvolver ao longo do tempo, a escolha de uma origem dos tempos não é importante.

Definição: um processo $Z = \{Z(t), t \in T\}$ diz-se estritamente estacionário se todas as distribuições finito-dimensionais permanecem as mesmas sob translações no tempo, ou seja,

$$F(z_1, \dots, z_n; t_1 + \Delta, \dots, t_n + \Delta) = F(z_1, \dots, z_n; t_1, \dots, t_n), \quad (4.2)$$

para quaisquer $t_1, \dots, t_n, \Delta$ de T .

Uma das consequências da equação (4.2) é poder assumir que, para séries estacionárias, todas as distribuições unidimensionais são fixas sob translações no tempo. Assim, dadas $\mu(t)$ a média e σ^2 a variância da série em questão, tem-se que

$$\mu(t) = \mu \quad (4.3)$$

$$\sigma^2(t) = \sigma^2 \quad (4.4)$$

para todo $t \in T$.

Note que, uma série pode ser estacionária apenas em períodos muito curtos, mudando de nível ou inclinação após certo tempo. Conforme Morettin e Tolo (2006), “a classe dos modelos ARIMA será capaz de descrever de maneira satisfatória séries estacionárias e séries não estacionárias, mas que não apresentem comportamento explosivo”. Para estes casos, pode ser tomado o intervalo de tempo que apresenta características estacionárias, desde que não muito curto, para que seja desenvolvida a análise.

Seguindo adiante, a normalidade de um processo é relacionada a distribuição de probabilidade de suas amostras. De acordo com Morettin e Tolo (2006), para uma série normal (ou gaussiana, aqui denominados sinônimos), “seu comportamento poderá ser descrito por um modelo linear”.

Definição: um processo $Z = \{Z(t), t \in T\}$ diz-se Gaussiano se, para qualquer conjunto $t_1, t_2, \dots, t_n$ de T , as variáveis $Z(t_1), Z(t_2), \dots, Z(t_n)$ têm distribuição normal n -variada.

Têm-se assim, que, para processos normais, o comportamento da série poderá ser determinado pelas médias e covariâncias.

Por fim, a independência de um processo, que diz respeito a existência ou não de relação entre os valores do processo, pode ser analisada, conforme Morettin e Tolo (2006),

através de funções de autocovariância (FACV). Os autores mostram que, para um processo estacionário real discreto, como o presente no caso prático a ser desenvolvido no presente trabalho de formatura, têm-se o seguinte.

Definição: Seja $Z = \{Z(t), t \in T\}$ um processo estacionário real discreto, de média zero, e função de auto covariância $\gamma_t = E\{Z_t Z_{t+\Delta}\}$, a FACV satisfaz as seguintes propriedades:

- 1) $\gamma_0 > 0$;
- 2) $\gamma_{-t} = \gamma_t$;
- 3) $|\gamma_t| \leq \gamma_0$;
- 4) γ_t é não negativa.

4.1.2 Modelos ARIMA

Conforme descrito anteriormente, a análise do presente projeto será feita por modelos ARIMA. Conforme Morettin e Tolo (2006), os modelos auto regressivos integrados e de média móvel (ARIMA) são uma classe de modelos paramétricos, em que a análise é feita no domínio do tempo. Alguns outros exemplos de modelos, também frequentemente usados em análises do tipo, são modelos de erro (ou de regressão), modelos auto regressivos e de médias móveis (ARMA), modelos de memória longa (ARFIMA), modelos estruturais e modelos não-lineares, todos estes últimos fora do escopo do presente projeto.

Os modelos ARIMA, de acordo com Morettin e Tolo (2006), são úteis para “descrever o comportamento de séries econômicas e sociais, onde os erros observados são auto correlacionados e influenciam a evolução do processo”. Três classes principais de processos podem ser descritas pelos modelos ARIMA, a primeira o caso do presente projeto:

- (i) processos lineares estacionários, em seus três sub-casos particulares: auto regressivos de ordem p , média móvel de ordem q , auto regressivos e de médias móveis de ordem p e q ;
- (ii) processos lineares não-estacionários homogêneos: generalização dos processos anteriores que supõem que o mecanismo gerador da série produz erros auto

correlacionados; geralmente, essas séries tornam-se estacionárias por meio de um número finito de diferenças;

- (iii) processos de memória longa: séries estacionárias que possuem uma função de auto correlação com decaimento muito lento, cuja análise necessitará de uma diferença fracionária.

Portanto, modelos auto regressivos integrados e de médias móveis de ordens p , d e q , denominados ARIMA(p , d , q), que tem ainda em sua generalização a inclusão de um operador sazonal, antes ainda não apresentado, são adequados para descrever todos os casos acima.

Detalhadamente, os três subcasos de processos lineares estacionários, a serem tratados no presente trabalho, estão definidos, conforme Morettin e Tolo (2006) nas equações (4.5), (4.6) e (4.7).

- a) Modelos auto regressivos (AR): pode ser definido como aquele que descreve um comportamento em que uma variável é uma combinação lineares dos valores obtidos anteriormente na série.

Definição: dado Z_t um processo estacionário, Φ os parâmetros auto regressivos de ordem p , e a_t um ruído branco, um modelo auto regressivo de ordem p , denotado por AR(p), é o mesmo que:

$$Z_t = \Phi_1 Z_{t-1} + \Phi_2 Z_{t-2} + \dots + \Phi_p Z_{t-p} + \dots + a_t \quad (4.5)$$

- b) Modelos de médias móveis (MA): pode ser definido como aquele que descreve um comportamento em que uma variável é uma combinação linear dos valores atuais e passados do termo denotado como ruído branco.

Definição: dado Z_t um processo estacionário, θ os parâmetros de médias móveis de ordem q , e a_t ruídos brancos, um modelo de médias móveis de ordem q , denotado por MA(q), é o mesmo que:

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \quad (4.6)$$

- c) Modelos auto regressivos e de médias móveis: naturalmente é a fusão dos dois modelos (AR e MA) já definidos. Isso permite, conforme Morettin e Tolo (2006) que se represente o processo com menor número de parâmetros em comparação a utilização dos modelos puros.

Definição: um modelo auto regressivo e de médias móveis de ordens p e q , denotado por ARMA(p, q), é o mesmo que:

$$Z_t = \Phi_1 Z_{t-1} + \cdots + \Phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \quad (4.7)$$

Para completar os modelos, a generalização para modelos auto regressivos integrados de médias móveis (ARIMA) permite que séries não estacionárias sejam também trabalhadas através de modelos ARMA(p, q). Isso é feito ao adicionar-se o parâmetro d na denotação já apresentada como ARIMA(p, d, q). Toma-se um número finito de diferenças d da série é possível torná-las estacionárias.

A denominação integrada deve-se ao fato da série original Z_t ser uma integral da série resultante W_t , já que esta última é formada a partir de d -ésimas diferenças da série original.

Box, Jenkins e Reinsel (1994) propuseram uma metodologia para auxiliar no processo de identificação do modelo ARIMA. Para isso, define-se os parâmetros para que melhor se ajustem ao conjunto de dados selecionados. A estratégia é baseada em um ciclo iterativo, no qual repete-se o ciclo caso o modelo não seja adequado. Recomenda-se selecionar mais de um modelo na primeira etapa a fim de evitar retrabalhos, e ao final, escolher entre os modelos aquele que melhor se ajustar aos dados. As etapas do processo estão apresentadas abaixo:

- 1) considerar uma classe geral de modelos para análise (denominada aqui como especificação);
- 2) identificar um modelo com base na análise de auto correlações, auto correlações parciais e outros critérios (denominada aqui como identificação);

- 3) estimar os parâmetros do modelo identificado (denominado aqui como estimação);
- 4) diagnóstico e verificação do modelo ajustado, através dos resíduos, a fim de verificar se o modelo está adequado para os fins desejados (denominado aqui como verificação).

4.1.3 Resíduos

Conforme metodologia apresentada no início do presente Capítulo, o último processo de análise de séries temporais é a verificação do modelo identificado, através da análise de seus resíduos, a ser apresentado neste Seção. A ideia é verificar se os parâmetros identificados no modelo ARIMA representam ou não adequadamente os dados. No mesmo processo, é possível perceber através de possíveis insuficiências uma sugestão de modelo alternativo.

Morettin e Toloi (2006) apresentam alguns testes para verificar a adequação do modelo principalmente através das auto correlações estimadas dos resíduos.

Definição: estimados Φ e θ e o modelo W_t , são chamados resíduos estimados, ou simplesmente resíduos as quantidades

$$\hat{a}_t = \theta^{-1}(B)\Phi(B)W_t \quad (4.8)$$

tal que B é operador definido através de um processo genérico X_t como

$$B^m X_t = X_{t-m} \quad (4.9)$$

e, portanto, $\Phi(B)$ e $\theta(B)$ são genericamente por $\emptyset$ como

$$\emptyset(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \dots - \emptyset_p B^p \quad (4.10)$$

Têm-se que, conforme Morettin e Toloi (2006), se o modelo for adequado, os $\hat{a}_t$ deverão estar próximos dos a_t e, portanto, deverão ser aproximadamente não-correlacionados, o que indica que se tratam de ruídos brancos.

4.2 Controle estatísticos de processo

Montgomery (2004) mostra que se há um processo que atende (e eventualmente supera) as expectativas de seus consumidores, ele deve ser o mais estável e replicável possível, para que seja operado com pequena variabilidade e qualidade alta do produto final. As sete principais ferramentas do Controle Estatístico de Processos (CEP), tratadas pelo autor, são poderosas no auxílio da resolução de problemas, alcançam estabilidade e aumentam a capacidade de redução da variabilidade de processos.

O resultado do modelo ARIMA de análise de séries temporais, mostrado anteriormente, tem também como resultado valores que em conjuntos podem ser tratados como um processo no tempo. Assim, da mesma forma que as ferramentas do CEP são utilizadas para processos industriais, é possível aplicá-las no meio financeiro através da análise das médias, variâncias e outros parâmetro para que se tenha uma noção se o objeto de estudo em questão está dentro do esperado ou não.

Conforme os conceitos de Montgomery (2004), o Controle Estatístico de Processos pode ser caracterizado como “um dos maiores desenvolvimentos tecnológicos do século XX porque é baseado em princípios confiáveis, fácil de usar, tem impacto significativo e pode ser aplicado a qualquer processo”. Suas sete principais ferramentas então são:

- 1) Histograma: gráfico que apresenta a frequência de distribuição dos dados do processo. Visa identificar padrões, dispersão e o tipo de distribuição estatística que melhor se encaixa para o processo;
- 2) Folha de controle: *checklist* com alguns itens pré-definidos que devem ser selecionados todas as vezes que eventos determinados acontecerem;
- 3) Gráfico de Pareto: gráfico de barras que coloca lado a lado em ordem decrescente a frequência de ocorrências das causas de não conformidade de um processo. Permite identificar quais são os maiores e menores causadores de impacto no processo e inclui uma curva de porcentagem acumulada;
- 4) Diagrama de causa e efeito: diagrama de Ishikawa ou diagrama espinha de peixe, busca estruturar as principais causas de um problema em questão. Ramifica-se também em sub-causas e reúne-se todas as possibilidades que podem contribuir para a ocorrência do problema tratado;

- 5) Diagrama da concentração de defeito: desenho técnico do produto final, mostra as vistas geométricas e busca entender se existe uma concentração das causas de não conformidade;
- 6) Diagrama de dispersão: gráfico da relação de duas (ou mais) variáveis que interferem no processo em questão. Busca-se uma correlação entre ambas e estabelece relações de causa e efeito;
- 7) Gráfico de controle: gráfico de agrupamento de valores observados no tempo. Utilizado como forma de acompanhamento e monitoramento do processo através de uma linha central e seus limites superior e inferior de controle.

Embora todas as ferramentas tenham a sua importância dentro da estrutura do CEP, no contexto do atual trabalho de formatura, a última delas, gráfico de controle, tem uma relevância maior, dada boa aderência para o controle de séries ao longo do tempo. Assim, em diante será desenvolvido com maiores detalhes as particularidades dessa ferramenta estatística, como será feita a seleção do gráfico de controle mais adequado ao modelo em questão e os detalhes específicos do desenvolvimento do gráfico escolhido.

4.2.1 Gráfico de controle

Conforme Montgomery (2004), “o gráfico de controle é uma ferramenta para descrever de maneira precisa exatamente o que é entendido por controle estatístico; como tal, ele pode ser usado em uma variedade de formas”. Desde aplicações on-line, coletando dados de alta frequência, até amostras de dados com um espaçamento maior de tempo, o gráfico de controle é capaz de monitorar e vigiar processos de forma a indicar se as respostas do processo estão sob controle ou não. É importante notar também, que pode ser de interesse tanto verificar se dados passados estavam sob controle como assegurar que os dados futuros indiquem um processo dentro dos limites de controle.

Assim, Montgomery (2004) define gráfico de controle como um *display* gráfico de uma variável estatística (W) que foi medida ou computada de uma amostra ao longo do tempo. O gráfico, conforme explicado, contém uma linha central (LC) definida a partir do valor médio (μ) da variável estatística. Além disso, o gráfico de controle contém outras duas linhas horizontais, calculadas a partir do desvio padrão (σ) da variável estatística, chamadas de limite superior de controle (LSC) e limite inferior de controle (LIC), que indicam pontos fora

de controle quando estiver fora da área delimitada superiormente pelo LSC e inferiormente pelo LIC.

Definição: dada uma amostra, uma variável estatística W cuja média e desvio padrão são respectivamente μ_W e σ_W e sendo L a distância entre a os limites de controle e a linha central (LC), genericamente, os limites de controle e a linha do gráfico de controle da estatística W são expressos como

$$LSC = \mu_W + L\sigma_W \quad (4.11)$$

$$LC = \mu_W \quad (4.12)$$

$$LIC = \mu_W - L\sigma_W \quad (4.13)$$

É importante destacar que, de acordo com Montgomery (2004), a maioria dos processos não é estatisticamente controlado e conseqüentemente podem estar sujeitos a variabilidades maiores. Os gráficos de controle podem ser ferramentas importantes para auxiliar na identificação das causas especiais, de forma a eliminá-las e reduzir a variabilidade dos processos.

Além disso, o autor cita ao menos cinco vantagens que fazem desta ferramenta do CEP ser tão popular:

- (i) Técnica comprovada para aumentar a produtividade;
- (ii) Prevenção de defeitos;
- (iii) Prevenção de ajustes desnecessários no processo;
- (iv) Informações diagnósticas sobre a amostra coletada;
- (v) Informações sobre a capacidade e estabilidade do processo.

Na prática, o gráfico de controle funciona como um teste de hipótese aplicado continuamente sobre as condições definidas em relação aos limites de controle, de forma que têm-se definidos:

H_0 : o processo está sob controle estatístico

H_1 : o processo não está sob controle estatístico

Sobre o processo em controle estatístico e a partir das hipóteses H_0 e H_1 , podem ser identificados e definidos os seguintes erros tipo I e tipo II. A razão dessa determinação é avaliar se o desempenho do gráfico de controle, determinado pela rapidez de detecção de um processo fora de controle, está de acordo com o desejado.

Definição: dado α a probabilidade de rejeitar-se a hipótese nula mesmo que verdadeira, o erro tipo I pode ser entendido como ter um processo sob controle, mas erroneamente ter um valor da estatística de controle W_t fora dos limites de controle. Conhecido como erro do “alarme falso”, têm-se definido

$$\alpha = P[W_t > LSC \mid \text{processo sob controle}] + P[W_t < LIC \mid \text{processo sob controle}] \quad (4.14)$$

Definição: dado β a probabilidade de aceitar-se a hipótese nula mesmo que falsa, o erro tipo II pode ser entendido como ter um processo fora de controle, mas erroneamente ter um valor da estatística de controle W_t dentro dos limites de controle. Conhecido como erro da “não-detecção”, têm-se definido

$$\beta = P[LIC < W_t < LSC \mid \text{processo fora de controle}] \quad (4.15)$$

Com o intuito de ter métricas definidas para comparar o desempenho de gráficos de controle, Montgomery (2004) define o *Average Run Length* (ARL), ou Comprimento Médio da Sequência, como o número médio de pontos da amostra que são marcados até que o gráfico indique uma condição fora de controle.

O ARL pode ser associado aos erros tipo I e tipo II, definidos nas equações (4.14) e (4.15), de forma que para um processo sobre controle, deseja-se ter o maior ARL possível (conhecido por ARL_0), enquanto que para um processo fora de controle, procura-se minimizar o ARL (conhecido como ARL_1). Têm-se então, que

$$ARL_0 = \frac{1}{\alpha} \quad (4.16)$$

$$ARL_1 = \frac{1}{1 - \beta} \quad (4.17)$$

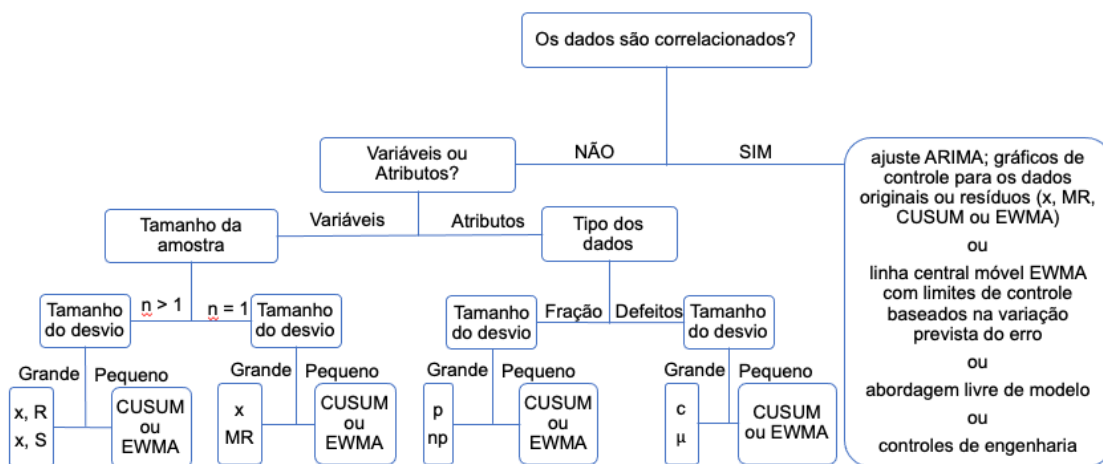
Em suma, um α muito reduzido (ARL_0 muito alto) pode levar na prática a um ARL_1 indesejavelmente alto (β também grande), e vice-versa. Isso acontece, pois com o distanciamento dos limites de controle da linha central (o que reduz os erros do tipo I), também é reduzida a capacidade do gráfico de detectar um processo que não está sob controle. Dessa forma, deseja-se balancear as duas probabilidades definidas de forma a ter um resultado estatisticamente relevante e com um desempenho dentro das expectativas.

4.2.2 Escolha do Gráfico de Controle

Montgomery (2004) mostra que, para a fórmula mostradas nas equações (4.11), (4.12) e (4.13), deriva-se alguns tipos de gráficos de controle, que podem ser diferenciados e selecionados através de alguns critérios, como a auto correlação dos dados, a classificação dos dados como variáveis ou atributos e a necessidade de observação de pequenas variações na curva ou não.

Nesse sentido, o autor desenvolveu um fluxograma para auxiliar na escolha do tipo de gráfico de controle mais adequado para o caso estudado, apresentado na Figura 4.1.

Figura 4.1 - Fluxograma de escolha do gráfico de controle



Fonte: Montgomery (2004)

Basicamente, o esquema apresenta um guia a ser seguido com perguntas que tem respostas objetivas a fim de filtrar caso a caso conforme suas particularidades. A primeira divisão do gráfico baseia-se na auto correlação dos dados, tem então duas alternativas:

- (i) em caso positivo, Montgomery (2004) assume que o tamanho da amostra é por intervalo de tempo e propõe então quatro alternativas: para ajustes ARIMA, aplicação de gráficos tradicionais seja para os dados originais ou para os resíduos; uso de gráfico de Média Móvel Exponencialmente Ponderada (EWMA) com linha central móvel; implementação de uma abordagem livre de modelo; e por último, eliminação das auto correlações (usualmente possível em processos industriais) usam-se controles de engenharia;
- (ii) em caso negativo, segue-se a segunda ramificação do fluxograma, que tem como próxima etapa definir se os dados são variáveis ou atributos, seguido de definição do tamanho da amostra para variáveis ou tipo dos dados para atributos. Por fim, busca-se a intenção de gráficos de detectarem um desvio grande ou pequeno na estatística observada.

Para o caso tratado no presente trabalho de formatura, assim como para boa parte das séries financeiras, há uma correlação significativa entre os dados a serem estudos (evidenciadas inclusive através da aplicação dos modelos de análise de séries temporais ARIMA). Assim, restam três opções a serem selecionadas e utilizadas para o modelo em questão:

- a) ajustes ARIMA, aplicação de gráficos tradicionais, como EWMA e *Cumulative Sum* (CUSUM), seja para os dados originais ou para os resíduos;
- b) gráfico de EWMA com linha central móvel;
- c) abordagem livre de modelo.

Devido à necessidade da identificação de pequenas mudanças no modelo e da facilidade de aplicação do gráfico EWMA, este foi o escolhido para representar o modelo do presente trabalho de formatura, portanto, daqui em diante, entende-se por gráfico de controle o gráfico EWMA (média móvel exponencialmente ponderada, em tradução direta para o português)

4.2.3 Gráfico de Controle EWMA

O gráfico de controle EWMA - *Exponentially Weighted Moving Average*, do inglês, ou Média Móvel Exponencialmente Ponderada, conforme Montgomery (2004), é frequentemente usado na modelagem e previsão de séries, já que o gráfico de controle pode ser visto como uma média ponderada de todos os eventos passados e presentes. Os principais parâmetros envolvidos na modelagem do EWMA são os valores de λ , constante que define o peso que é dado aos valores passados, e o valor do limite de controle L, que pode ser determinado através de simulação.

Definição: dados $0 < \lambda \leq 1$ uma constante, W_t o valor da variável estatística e o valor inicial z_0 , necessário para o a primeira amostra, assumido como $z_0 = \mu_0$, a média móvel exponencialmente ponderada é definida como

$$z_i = \lambda W_i + (1 - \lambda)z_{i-1} \quad (4.18)$$

Genericamente, realiza-se as devidas substituições na equação (4.18) dos valores de $z_{i,j}$ para $j = 2, 3, \dots, t$ e tem-se que

$$z_i = \lambda \sum_{j=0}^{i-1} (1 - \lambda)^j W_{i-j} + (1 - \lambda)^i z_0, \quad (4.19)$$

e o fator $\lambda(1 - \lambda)^i$ decai geometricamente com tempo da amostra. Conforme Montgomery (2004), “devido a queda geométrica dos pesos quando conectados pela curva suavizada, o EWMA é chamado também média móvel geométrica”.

Assim, o gráfico de controle EWMA pode ser construído através das amostras z_i *versus* o número da amostra ou tempo i . Os limites de controle do gráfico, são calculados conforme as equações (4.20), (4.21) e (4.22).

$$LSC = \mu_0 + L\sigma \sqrt{\frac{\lambda}{(2 - \lambda)} [1 - (1 - \lambda)^{2i}]} \quad (4.20)$$

$$LC = \mu_0 \quad (4.21)$$

$$LSC = \mu_0 - L\sigma \sqrt{\frac{\lambda}{(2 - \lambda)} [1 - (1 - \lambda)^{2i}]} \quad (4.22)$$

As determinações dos parâmetros L e λ , no atual caso deste trabalho de formatura, serão feitas através de simulações para que o ARL_0 e o ARL_1 do gráfico estejam de acordo com o esperado e o desempenho do gráfico seja aceitável.

5 APLICAÇÃO PRÁTICA

A partir das referências teóricas apresentadas nos Capítulos 2, 3 e 4, o capítulo atual do presente Trabalho de Formatura tem como objetivo aplicar os conceitos no caso prático e real sobre os resultados mensais do IPCA, índice de inflação do Brasil. Os detalhes da metodologia de aplicação prática que ainda não foram demonstrados anteriormente também serão apresentados neste capítulo.

É importante destacar que todos os códigos de programação e simulação foram feitos na linguagem de programação R que utiliza dos pacotes de linguagem de programação necessários para tais aplicações. O programa completo está apresenta ao final deste trabalho de formatura, no Apêndice A.

5.1 Considerações iniciais

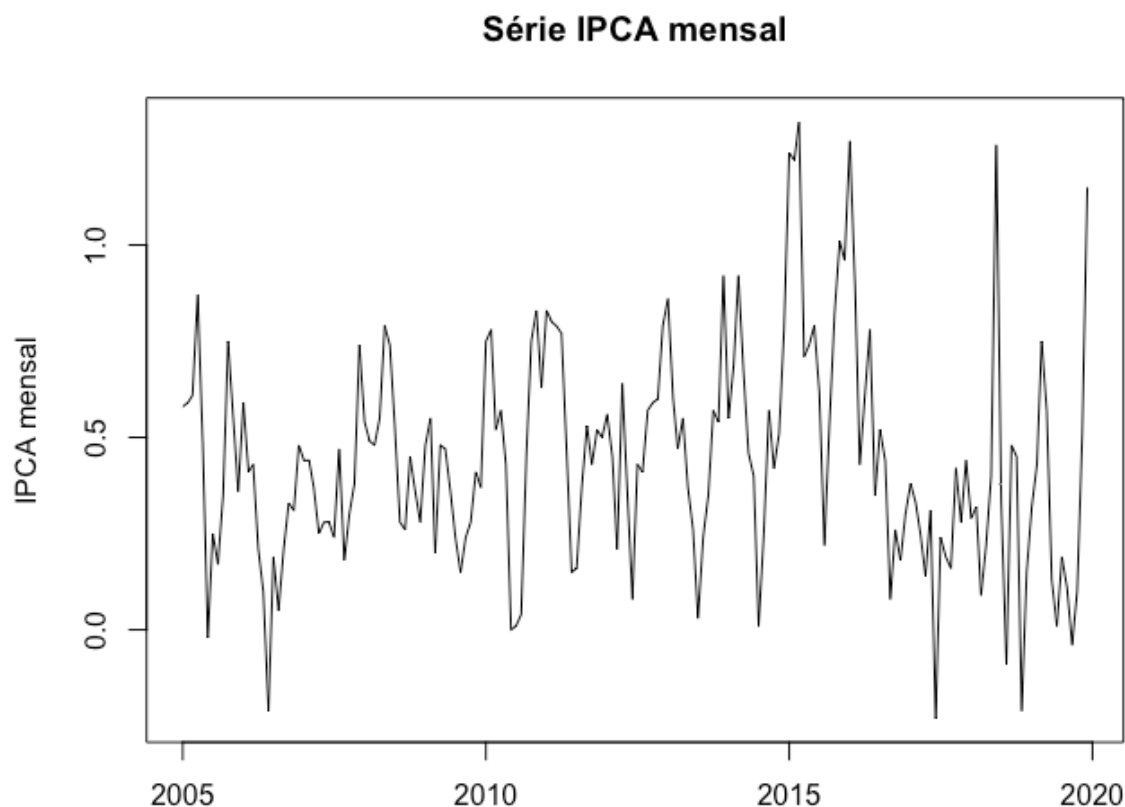
Conforme explicado nos capítulos anteriores, a série a ser trabalhada é o IPCA mensal, considerado aqui entre o período de janeiro de 2005 e dezembro de 2019. Os dados podem ser obtidos em fonte pública em bases de dados do IBGE, instituição responsável pelo seu cálculo e divulgação. Na Tabela 5.1 estão apresentadas algumas estatísticas apenas descritivas sobre a série inicial. Além disso, Figura 5.1 está o gráfico que mostra o comportamento da série ao longo do tempo.

Tabela 5.1 – Estatísticas descritivas de série inicial

IPCA Mensal	
Início	2005-01
Final	2019-12
Frequência	Mensal
Observações	180
Média	0.004439
Variância	0.000008130
Desvio Padrão	0.002851
Máximo	0.01320
Mínimo	-0.002300
Mediana	0.004301

Fonte: Elaborada pelo autor (2021)

Figura 5.1 - Série IPCA mensal



Fonte: Elaborada pelo autor (2021)

Nota-se que a série temporal do IPCA mensal tem um comportamento que se assemelha ao de uma série estacionária. Dessa forma, não serão feitas transformações iniciais e realiza-se a aplicação da metodologia ARIMA diretamente na série original.

5.2 Preparativos para aplicação do modelo

A fim de aplicar-se um modelo ARIMA, antes mesmo da definição do modelo, conforme explicado anteriormente, deve-se verificar, em relação a série temporal do IPCA, as questões de estacionariedade, normalidade e independência. Os testes necessários para cada uma das propriedades estão apresentados nesta Seção.

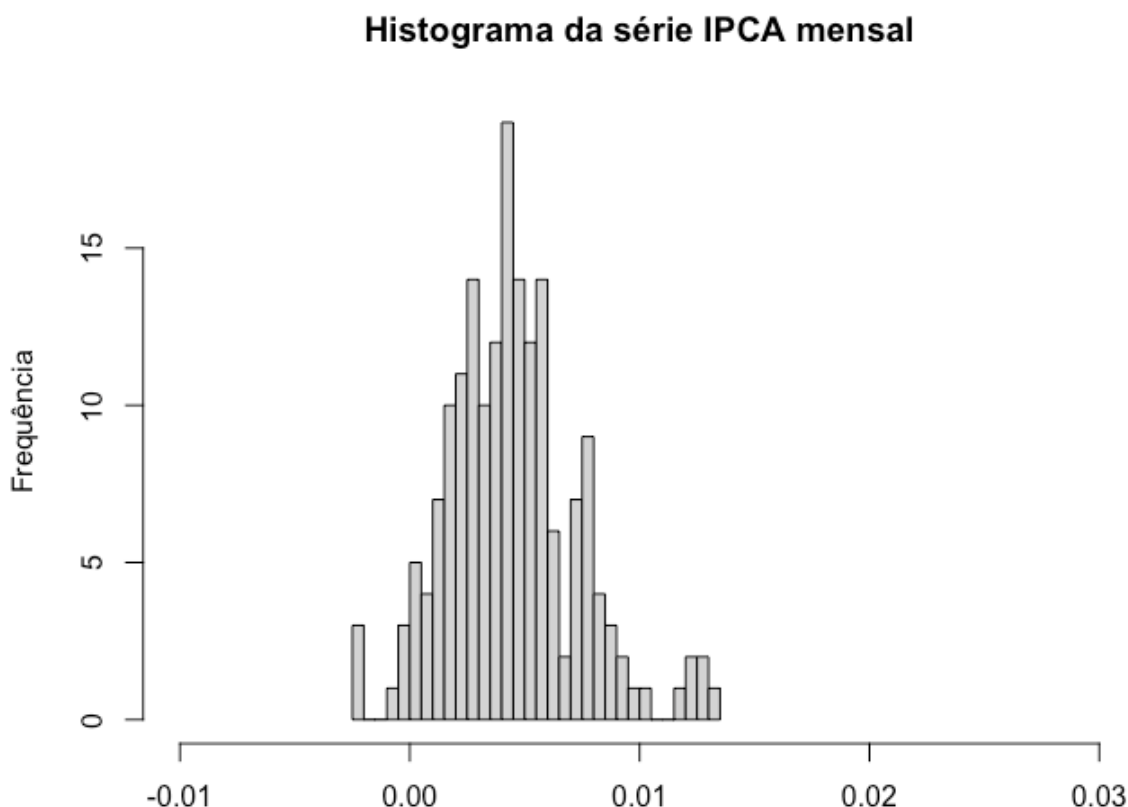
- a) Estacionariedade: a estacionariedade da série é testada através do teste *Augmented Dickey-Fuller* (ADF). Em suma, testa-se a hipótese nula de não-estacionariedade da série. Em caso de rejeição, aceita-se uma hipótese alternativa de série estacionária. Em caso de aceite da hipótese nula, Morettin e Toloí (2006) sugerem tomar diferenças da série até que a hipótese de não-estacionariedade seja rejeitada. Para a série IPCA mensal trabalhada, têm-se uma estatística calculada de -4.7581 e um $p\text{-value} < 0.01$.

Tal resultado vai de acordo com as expectativas e confirma a hipótese alternativa de estacionariedade da série.

- b) Normalidade: a normalidade de série será testada de três formas diferentes: primeiro, será analisado um histograma dos dados, que observa se a distribuição está adequada a forma padrão de uma curva normal; após isso, será construído um *Q-Q Plot* para os dados, com o intuito de observar como os quantis da distribuição se comportam quando comparados com os quantis teóricos da distribuição normal; por fim, será realizado o teste de *Shapiro-Wilk*, que busca confirmar estatisticamente se os dados seguem uma distribuição normal.

O histograma para a série IPCA mensal está apresentado na Figura 5.2. Percebe-se que os dados não têm uma adesão perfeita a uma distribuição normal, porém assemelham-se de forma geral, com um desvio maior nas pontas do histograma, que concentram mais valores que o esperado.

Figura 5.2 - Histograma da série IPCA mensal

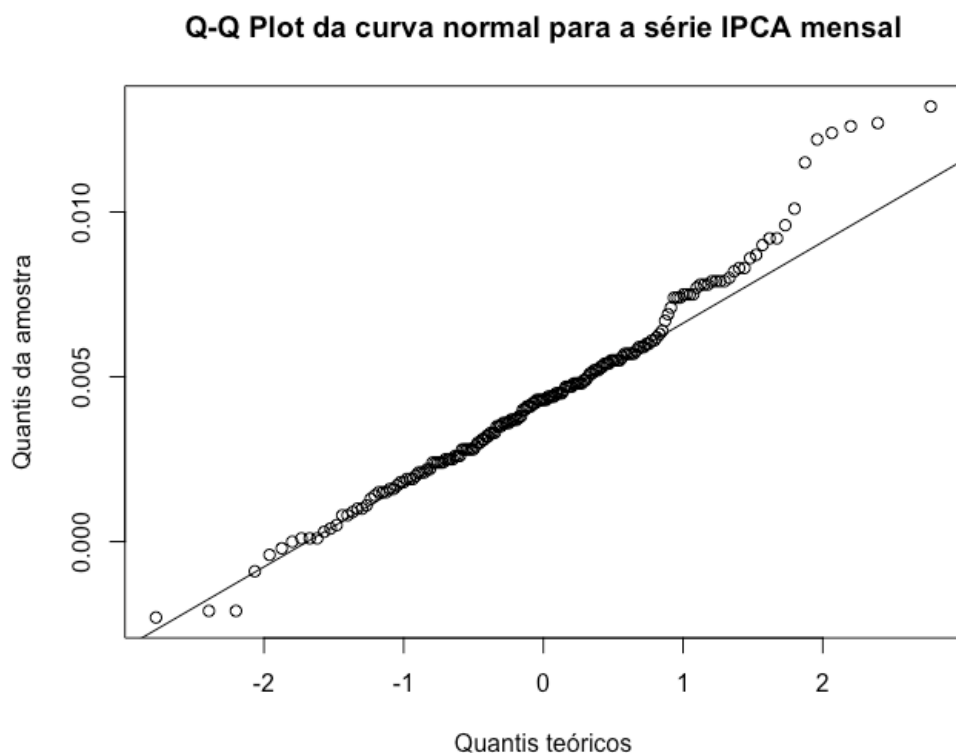


Fonte: Elaborado pelo autor (2021)

Em relação ao *Q-Q Plot* para a série em questão, mostrado na Figura 5.3, verifica-se a mesma tendência percebida no histograma: há uma boa adesão dos pontos da série em boa

parte da reta normal, mas nos extremos, principalmente o quantil superior, há um desvio mais significativo dos dados.

Figura 5.3 - Q-Q Plot normal para a série IPCA mensal



Fonte: Elaborado pelo autor (2021)

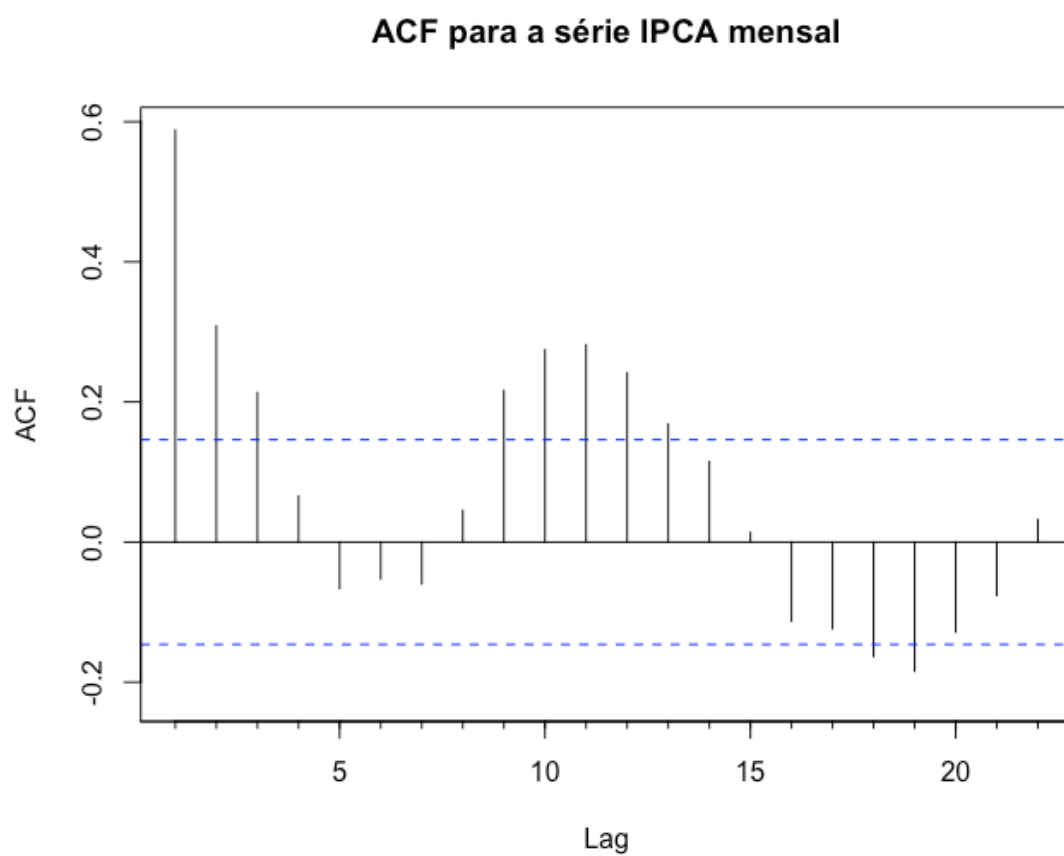
Por fim, o teste de normalidade de *Shapiro-Wilk* mostrou que um *p-value* de 0.003248 e valor da estatística de 0.97587. Como o *p-value* é menor que 0.05, valor geralmente utilizada para aceitação da hipótese nula da normalidade, rejeita-se a hipótese nula e assim não se pode afirmar que os dados seguem uma distribuição normal. Ainda assim, Kumiega e Van Vliet (2012), mostram que, ainda que existam assimetrias em relação às series financeiras e à curva normal, os modelos possuem resultados satisfatórios e confiáveis.

Portanto, apesar da não confirmação da normalidade da série IPCA mensal, os testes ainda assim mostraram que há uma tendência normal a ser seguida pela série, de forma que se retirados os pontos extremos, ter-se-ia a hipótese de normalidade de *Shapiro-Wilk* aceita. Assim, assume-se uma série adequada quanto a esse critério.

- c) Independência: por parte da independência da série IPCA mensal, busca verificar a mesma através das funções de auto correlação (ACF) e auto correlação parcial (PACF)

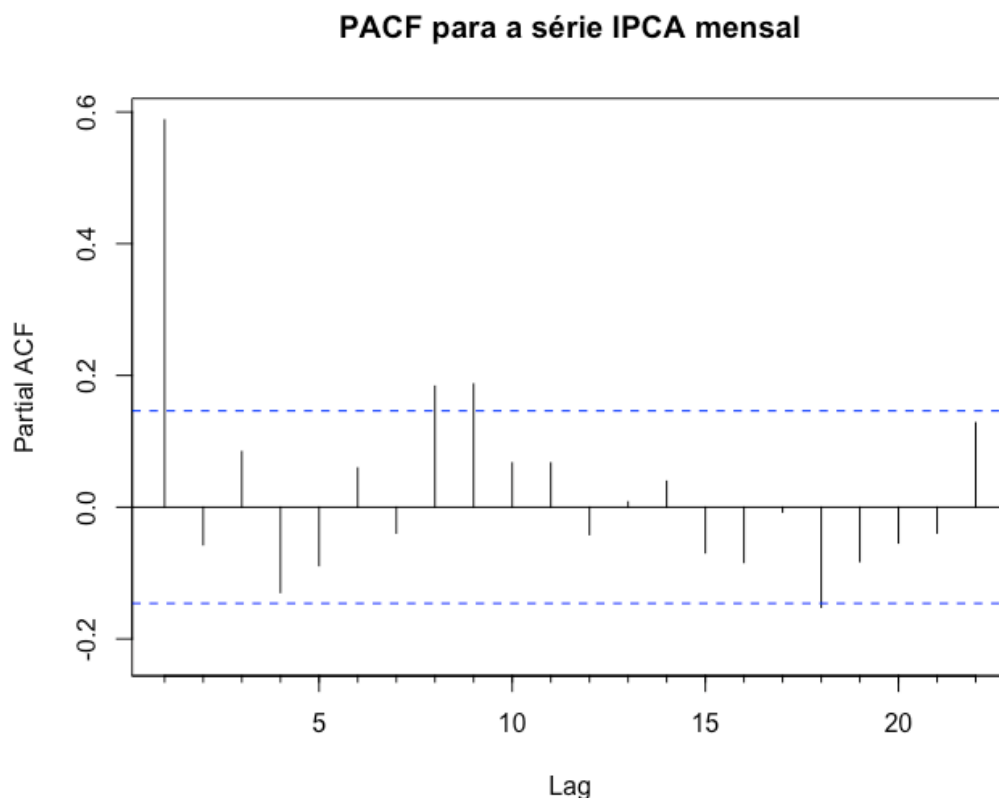
dos dados e define-se também prováveis parâmetros do modelo. As funções estão desenvolvidas na Figura 5.4 e Figura 5.5, respectivamente.

Figura 5.4 - ACF da série IPCA mensal



Fonte: Elaborada pelo autor (2021)

Figura 5.5 - PACF da série IPCA mensal



Fonte: Elaborada pelo autor (2021)

Verifica-se pelas funções que é uma série com auto correlações, principalmente nos primeiros *lags*, em que os valores críticos se apresentam muito acima da linha pontilhada azul das funções.

Conclui-se aqui, portanto, que após a avaliação dos critérios de estacionariedade, normalidade e independência, a série IPCA mensal apresenta um comportamento suficiente para que seja aplicada dentro de um dos modelos ARIMA.

5.3 Modelo ARIMA

Dadas as confirmações realizadas na Seção 5.2, busca-se agora definir o modelo ARIMA(p,d,q) que melhor ajusta a série em questão. Sabe-se, a partir dos testes de independência dos dados da série realizados, que os primeiros lags terão uma significância no modelo. Busca-se também saber, por outro lado, se há parâmetro de médias móveis que satisfazem tais necessidades. Além disso, o modelo deve ser capaz de identificar se há grau relevante de sazonalidade que deverá ser também implementada na série.

Para isso, utilizou-se método de simulação de modelos ARIMA disponível para a linguagem de programação R. A função realiza uma varredura de diversas combinações de (p,d,q) com o intuito de verificar quais são os parâmetros mais relevantes para a série em questão. No fim, após várias rodadas, escolheu-se o modelo ARIMA(2,0,3) ou ARMA(2,3) com média diferente de zero, cujas estimativas dos parâmetros do modelo com os respectivos erros padrão estão apresentados na Tabela 5.2.

Tabela 5.2 - Coeficientes do modelo ARMA(2,3) com média diferente de zero

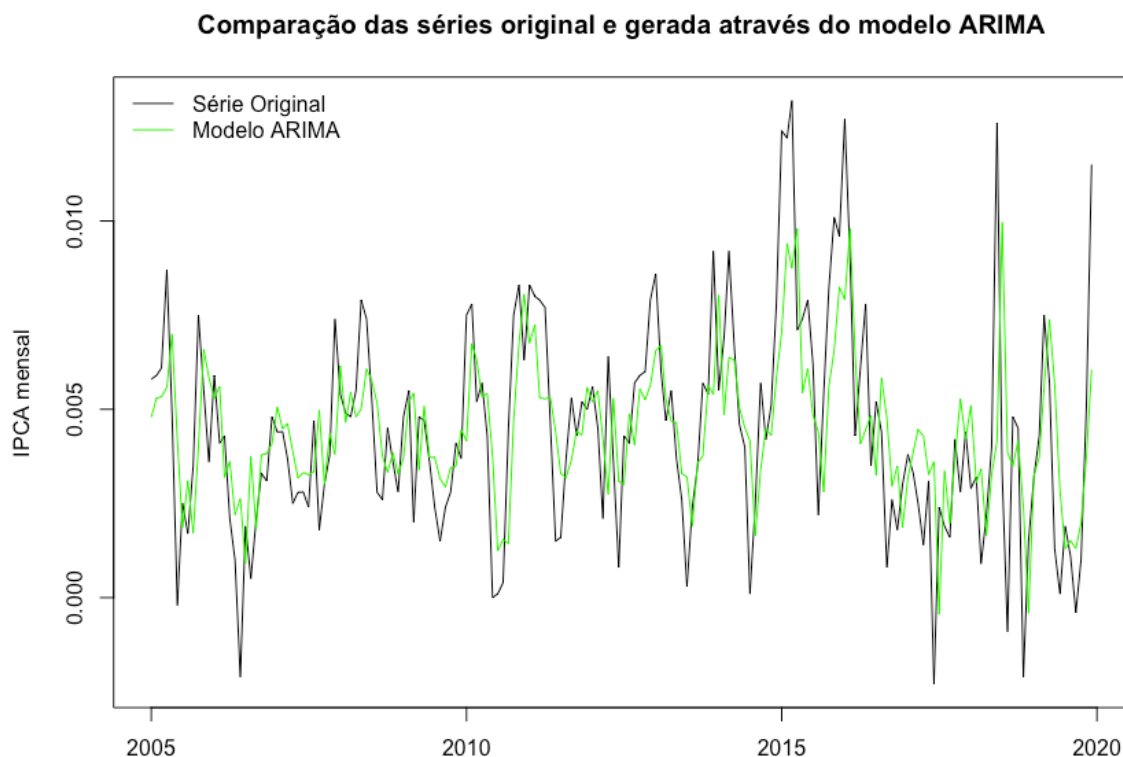
Coeficientes	Valor	Desvio Padrão	$ \frac{Valor}{Desvio\ Padrão} $
ar1	1.3242	0.1025	12.922580
ar2	-0.7691	0.0927	8.298883
ma1	-0.7163	0.1035	6.917553
ma2	0.2491	0.0894	2.785466
ma3	0.4126	0.0849	4.862381
média	0.0045	0.0004	12.643092

Fonte: Elaborado pelo autor (2021)

Assim, têm-se ajustada a série com o modelo ARMA(2,3) com média diferente de zero para o IPCA mensal. Nota-se que, conforme esperado, o primeiro lag em relação a data atual tem correlação positiva e mais relevante do que todos os outros parâmetros do modelo.

Figura 5.6 apresenta o gráfico que contém a série original e a série gerada através do modelo. Nota-se a boa adesão dos sinais gerados através do modelo ARMA(2,3) com média diferente de zero e a série original do IPCA mensal. É perceptível, em um primeiro momento, que o modelo tem uma melhor aderência com os dados que estão mais próximos a média e abre um gap maior para os dados que estão mais afastados do centro do gráfico.

Figura 5.6 - Séries original e modelo ARMA(2,3) com média diferente de zero



Fonte: Elaborado pelo autor (2021)

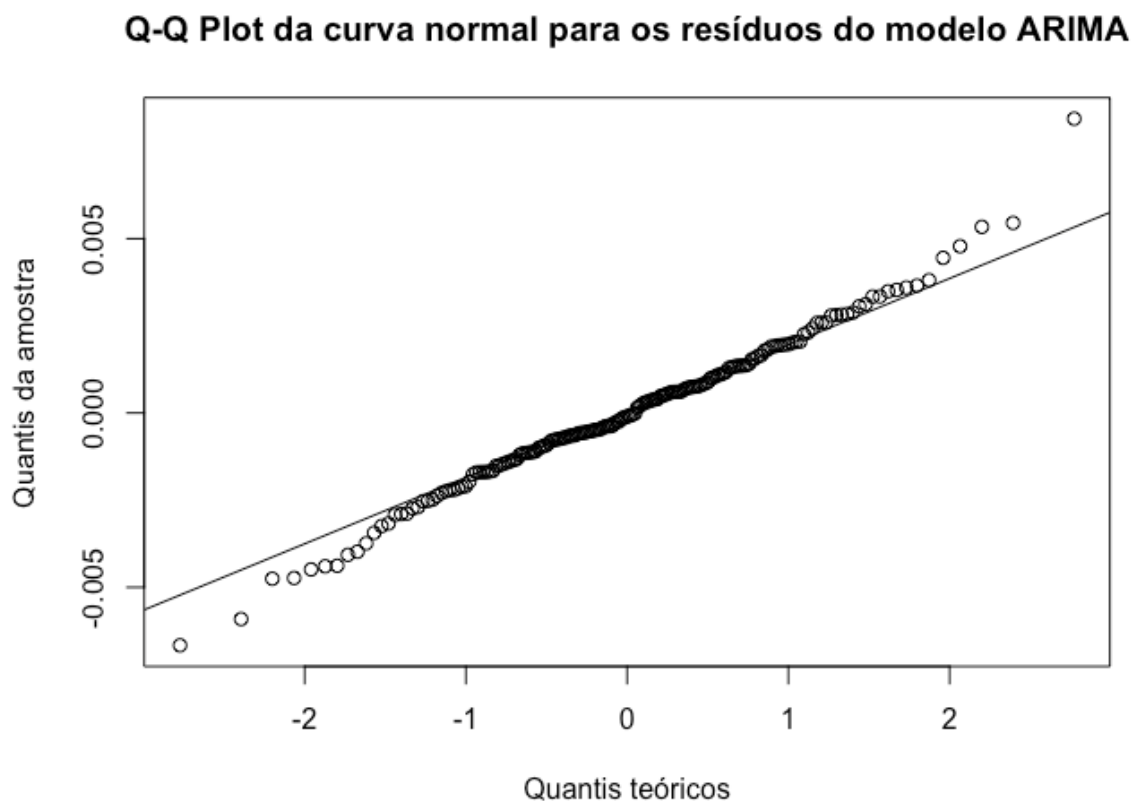
5.4 Análise dos resíduos

Conforme especificado no Capítulo 4, o passo final para validação do modelo ARMA(2,3) com média diferente de zero sobre o IPCA mensal é verificar se os resíduos tratam-se de ruídos brancos. Isto é, verificar se os resíduos se encaixam nos critérios de normalidade e independência. Para isso, serão realizados os testes *Q-Q Plot* e Shapiro-Wilk para a normalidade e as funções de auto correlação e auto correlação parcial assim como o teste de Box-Pierce-Ljung.

- a) Normalidade: assim como foi feito para os dados originais da série IPCA mensal, a normalidade dos resíduos gerados pelo modelo aplicado será testada através de um *Q-Q Plot* e da aplicação do teste estatístico de normalidade Shapiro-Wilk.

O *Q-Q Plot*, mostrado na Figura 5.7, mostra uma boa aderência dos resíduos a curva normal, com pequenos desvios apenas no quantis extremos, como já havia sido observado para os dados originais, porém de menor intensidade e volume para os resíduos.

Figura 5.7 - Q-Q Plot normal para os resíduos do modelo ARMA(2,3) com média diferente de zero



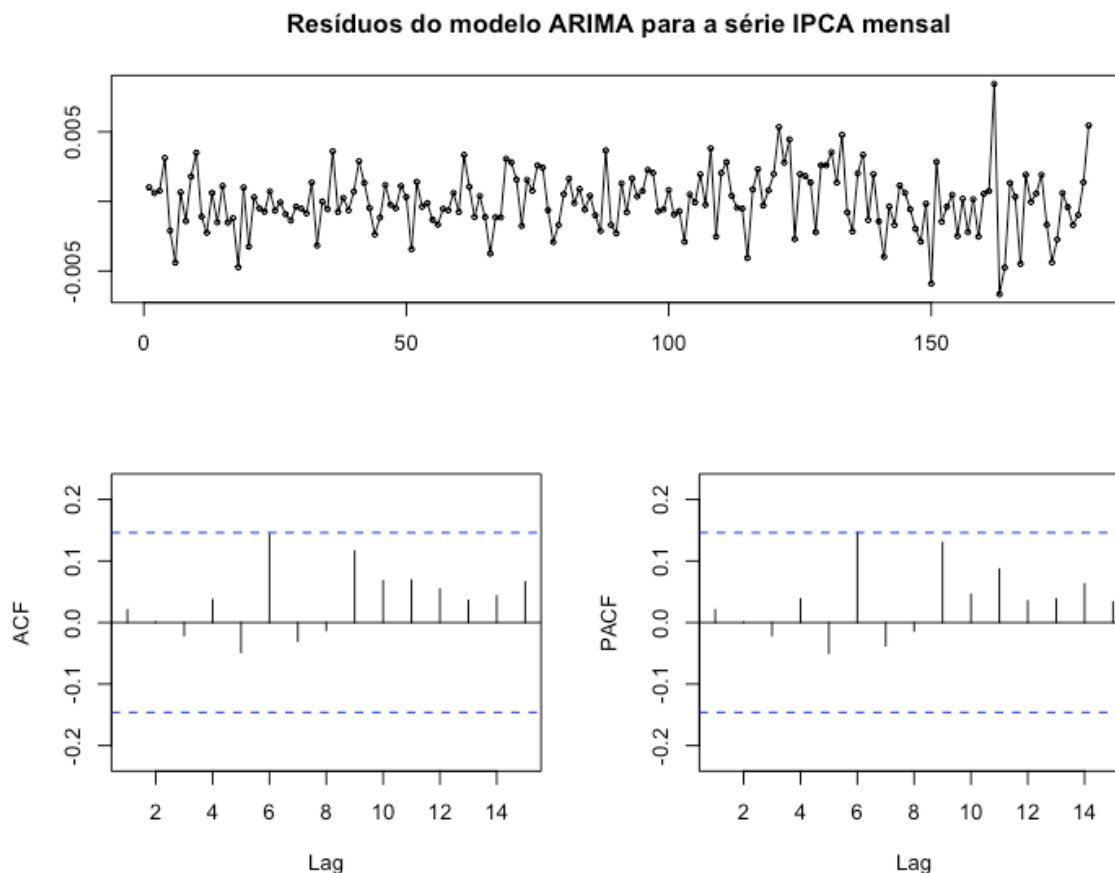
Fonte: Elaborado pelo autor (2021)

O teste de normalidade de Shapiro-Wilk, ao contrário do que foi apresentado para os dados da série original, mostra uma ótima aderência dos resíduos à curva normal e confirma a adequação deles ao critério de normalidade. Foi observada uma estatística W de 0.98878 e um p -value de 0.1659, bem acima do mínimo necessário de 0.05.

- b) Independência: da mesma forma como foi testado para os dados originais da série mensal do IPCA, a independência dos resíduos do modelo será testada através das funções de auto correlação e auto correlação parcial, busca-se confirmar que não há valores acima dos níveis críticos. Além disso, será realizada o teste estatístico de Box-Pierce-Ljung.

As funções de correlação e auto correlação parcial, assim como a disposição dos resíduos no tempo, estão apresentados na Figura 5.8. Nota-se que em ambos os gráficos de função de auto correlação não há *lags* que passam dos valores críticos, pode-se aceitar a independência dos resíduos por esse primeiro critério.

Figura 5.8 - Análise ACF e PACF para os resíduos do modelo

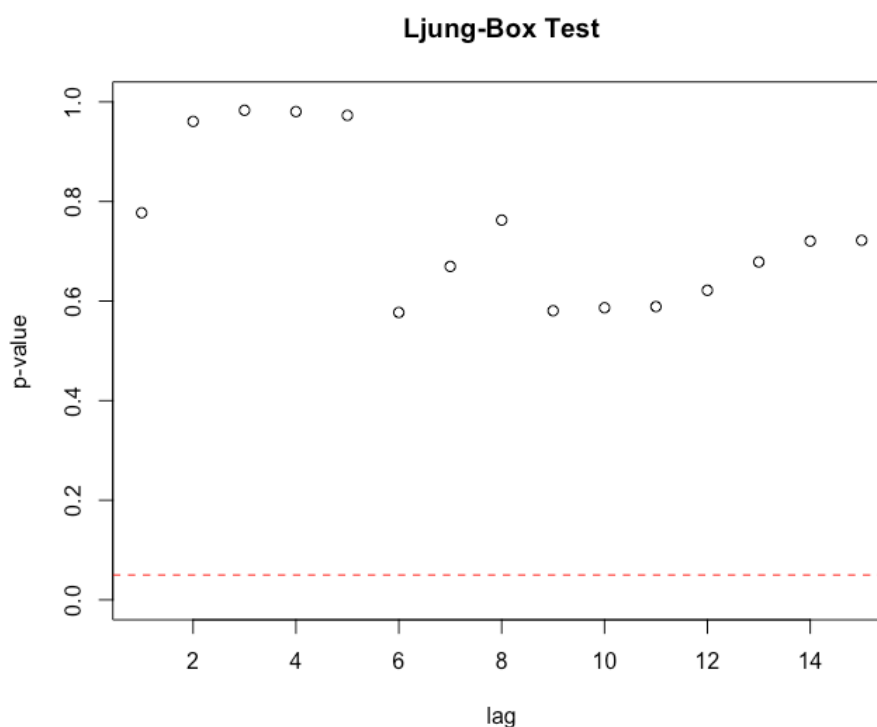


Fonte: Elaborado pelo autor (2021)

Além disso, Box e Pierce (1970) criaram um teste para testar se os resíduos de um modelo de fato podem ser assumidos como ruídos brancos. A hipótese nula do teste diz que os resíduos são independentes e identicamente distribuídos. Assim, busca-se aceitar o H_0 para que a validade do modelo seja finalmente confirmada.

Conforme a Figura 5.9, para todos os lags avaliados pelo teste, o mínimo *p-value* foi por volta de 50% e leva a aceitação de H_0 devido a enorme diferença que existe entre o mínimo *p-value* de ao menos 5% que é estabelecido pelo teste.

Figura 5.9 – Teste Ljung-Box para os resíduos do modelo ARIMA



Fonte: Elaborado pelo autor (2021)

5.5 Controle do modelo e geração dos sinais

Com a validade do modelo finalmente confirmada e os dados já gerados para a toda a extensão do período, o próximo passo é, a partir dos resultados do modelo, gerar um gráfico de controle do tipo EWMA para a série simulada do IPCA mensal.

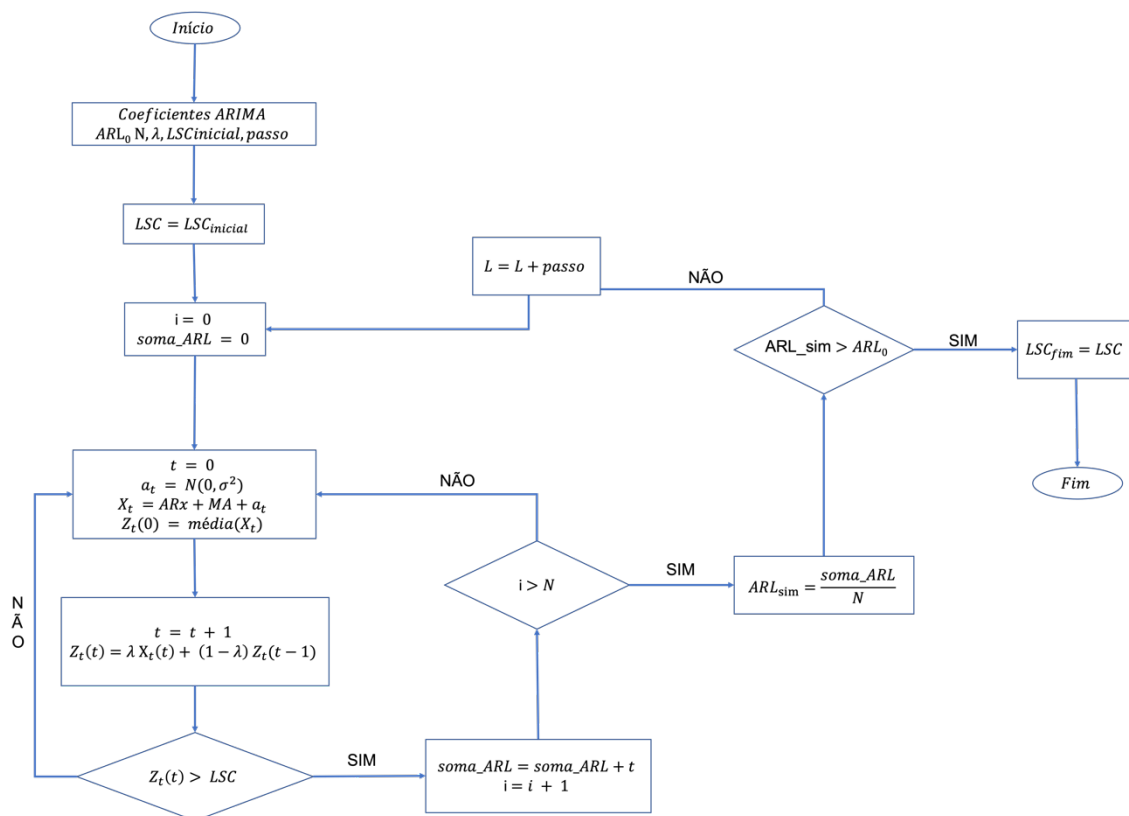
O primeiro passo para a determinação do gráfico de controle é a definição dos limites de controle. É importante notar aqui que, devido a natureza e o histórico do índice de inflação IPCA, há um interesse maior em determinar valores que estão acima de um limite de controle superior e tem-se assim que um limite inferior de controle possui uma relevância muito menor. Dessa forma, as simulações a partir desse momento terão como objetivo determinar apenas um limite superior de controle para o modelo gerado.

O cálculo deles é feito através de simulação, que leva em conta definições iniciais de um ARL_0 e de um constante lambda, os dois principais parâmetros do gráfico de controle EWMA. Além disso, deve-se especificar também N, o limite de controle inicial, o passo de atualização do limite de controle e os valores dos coeficientes determinados no modelo ARMA(2,3) com média diferente de zero. Os valores iniciais empregados na simulação estão apresentados a seguir e ilustrados na Figura 5.10.

Vale ressaltar que foram testados diversos valores de lambda, obtendo resultados consistentes no modelo apenas para o valor definido para a constante.

- ARL_0 : 18 meses -
- Lambda (λ): 1
- N: 1000 – número de simulações feitas pelo algoritmo;
- LSC inicial: 0.001 – limite de controle que será variado conforme o passo a cada simulação até que o ARL testado ultrapasse o ARL_0 ;
- Passo: 0.00001 – valor que será aumentado do LC inicial a cada simulação;
- Coeficientes do modelo ($ar1$, $ar2$, $ma1$, $ma2$, $ma3$, média, desvio padrão): parâmetros do modelo ARMA(2,3) desenvolvido para que seja considerados nas simulações.

Figura 5.10 - Fluxograma para simulação e definição dos limites de controle

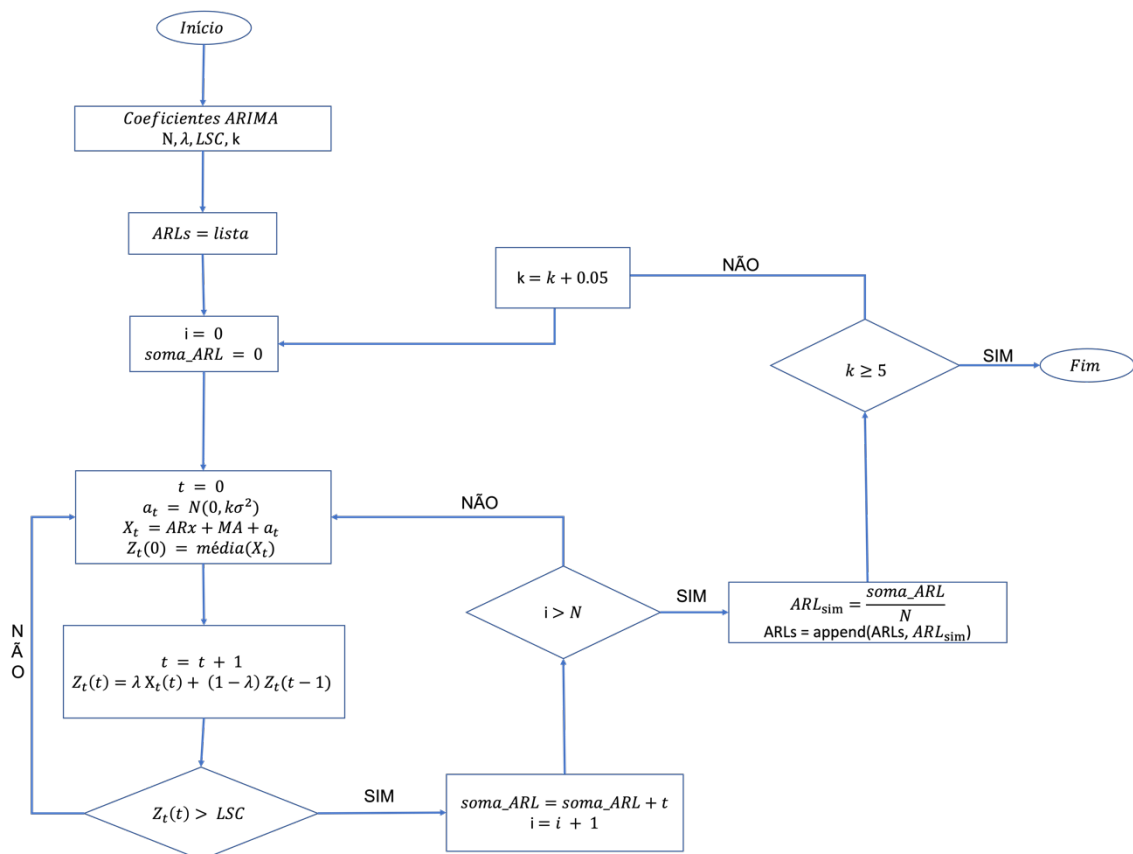


Fonte: Elaborado pelo autor (2021)

Conforme as premissas indicadas, ao final da simulação, chegou a um valor de limite superior de controle igual a 0.00842. Isso indica, em um primeiro momento, um valor de LSC 0,0039 acima da linha central do gráfico. É esperado, conforme a simulação, que existe um alarme falso a cada 18 meses, de acordo com o esperado inicialmente.

Após a definição do LSC, o próximo passo é determinar como a variação do desvio padrão dos resíduos afeta o valor de ARL encontrado anteriormente. Para isso, é feita outra simulação com os dados que utiliza dos mesmos parâmetros anteriormente definidos, com exceção de duas modificações principais: o LSC é fixado no valor encontrado anteriormente e o valor que se deseja variar é o desvio padrão dos resíduos do modelo, faz-se através da multiplicação por um valor k . Define-se então um $k=1$ para início de simulação e um passo de 0.05 para incremento a cada N simulação, deixa-se variar até k igual a 5, o que totaliza 80 valores testados. O passo a passo da simulação para determinação do ARL para os dados fora de controle está descrito no fluxograma da Figura 5.11.

Figura 5.11 - Fluxograma da simulação do ARL



Fonte: Elaborado pelo autor (2021)

A simulação mostrou que há uma sensibilidade maior para um incremento de até duas vezes no valor do desvio padrão, dado que há uma redução de quase dois terços do valor inicial (de 18 para 6.22). Para variação ainda maiores, há uma variação menos acentuada, de forma que há uma estabilização em um ARL próximo de 4 para grandes valores de k . Isso mostra que, o gráfico tem capacidade de detectar variações no processo com eficácia e encontra-se em um nível adequado, já que para valores expressivamente menores do desvio

padrão dos resíduos, têm-se naturalmente valores de ARL muito menores, o que impactaria negativamente o desempenho do gráfico.

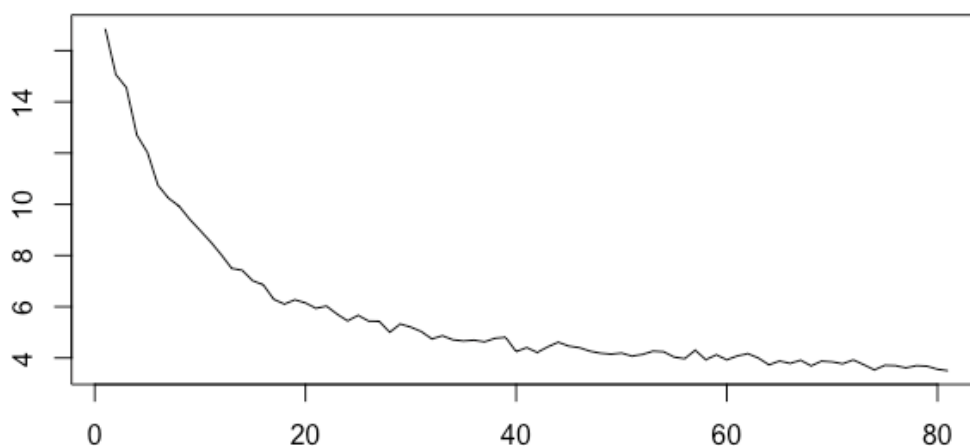
A Tabela 5.3 resume dos principais valores de k para os valores de ARL_1 . Além disso, a Figura 5.12 mostra o gráfico com os valores simulados para cada um dos k e mostra a forte sensibilidade no início do aumento do desvio padrão também está apresentado a seguir.

Tabela 5.3 - Valores de k e ARL

k	ARL_1
1	18
1,5	8,21
2	6.22
3	4.27
4	3.79
5	3.75

Fonte: Elaborado pelo autor (2021)

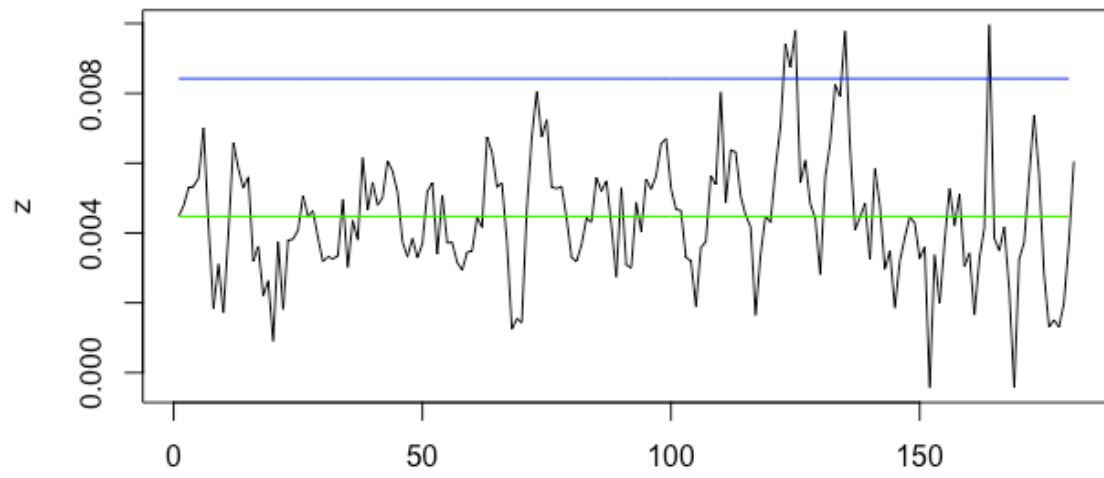
Figura 5.12 - Valores do ARL para variação do k



Fonte: Elaborado pelo autor (2021)

Por fim, apresenta-se na Figura 5.13 o gráfico de controle final obtido após as simulações e a determinação do limite superior de controle. O limite superior de controle, representado pela linha azul contínua em $z = 0,00842$ no gráfico faz com que 5 valores da amostra total de 180 valores do modelo estejam identificados como fora de controle.

Figura 5.13 - Gráfico de controle para a série ARIMA



Fonte: Elaborado pelo autor (2021)

6 ANÁLISE DOS RESULTADOS

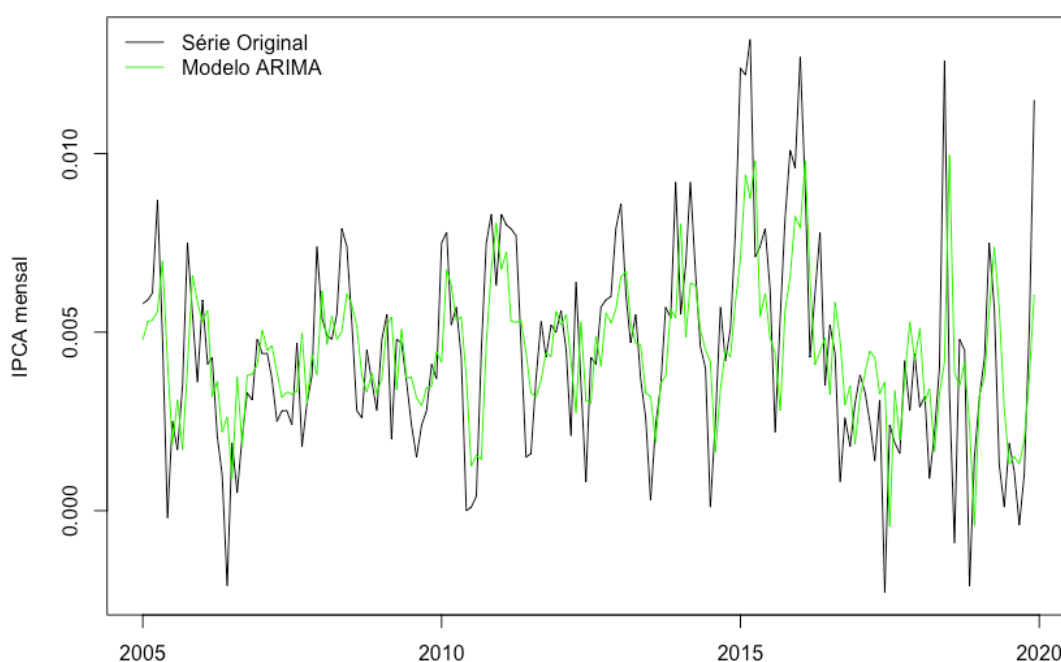
No capítulo anterior, foram aplicados os conceitos teóricos sobre análise de séries temporais e gráficos de controle mostrados no quarto capítulo sobre a série de dados IPCA mensal apresenta no segundo capítulo do presente Trabalho de Formatura. Os resultados obtidos tanto para o modelo ARMA(2,3) com média diferente de zero como para a aplicação do resultado do modelo sobre o gráfico de controle EWMA serão explorados nesse presente capítulo, com o objetivo de mostrar consequências diretas e de testar a eficácia dos mesmos.

6.1 Modelo ARIMA

Na Figura 6.1 está apresentado novamente o gráfico que mostrou o resultado da comparação entre a série original IPCA mensal e os sinais gerados após a aplicação da série original sobre o modelo ARMA(2,3) com média diferente de zero.

Figura 6.1 - Séries original e modelo ARMA(2,3) com média diferente de 0

Comparação das séries original e gerada através do modelo ARIMA



Fonte: Elaborado pelo autor (2021)

Nota-se que, há duas análises diretas para serem feitas a partir do resultado obtido: primeiro, ao analisar a eficácia do gráfico, busca-se verificar estatisticamente qual a diferença entre os valores reais e os sinais gerados através do modelo ARIMA; além disso, é possível realizar uma projeção dos valores do modelo para datas futuras não abordadas na análise e

verificar como será o desempenho frente aos dados reais do IPCA. Essas duas análises serão discorridas a seguir.

6.1.1 Comparação com as expectativas do mercado

Com o intuito de analisar a precisão dos resultados gerados pelo modelo ARMA(2,3) com média diferente de zero quando aplicado sobre a série IPCA mensal, foram analisadas para cada um dos meses do período o módulo da diferença entre os dois dados. Os principais resultados obtidos para a série mensal do módulo das diferenças estão apresentados na Tabela 6.1.

Na Tabela 6.2 estão apresentados os resultados para a mesma análise da Tabela 6.1, mas desta vez compara o módulo das diferenças do IPCA mensal real e da mediana de previsão do mercado inteiro para o índice, que pode ser extraída através dos dados do relatório Focus, disponível na base de dados do BCB e atualizado todas as segundas-feiras.

Tabela 6.1 – Estatísticas da diferença entre as séries

Módulo das diferenças IPCA mensal real e modelo ARIMA	
Média	0.0016
Mediana	0.0013
Variância	0.0000019
Desvio Padrão	0.0014
Máximo	0.0084
Mínimo	0.000012

Fonte: Elaborada pelo autor (2021)

Conforme os resultados da ambas as tabelas, vê-se que a média da diferença das previsões do mercado e da expectativa do modelo são consideráveis, mostra que o modelo ARMA(2,3) com média diferente de zero, quando aplicado para o IPCA mensal para os períodos em questão, apresenta resultados não tão satisfatórios se comparado aos modelos de expectativa usados pelo mercado atualmente. Isso é explicado, ainda conforme comparação das Tabela 6.1 e Tabela 6.2, pela grande diferença entre as máximas diferenças encontradas, que evidencia a dificuldade do modelo em adaptar-se a choques inflacionários inesperados.

Tabela 6.2 – Estatísticas da diferença entre as séries

Módulo das diferenças IPCA mensal real e expectativa do mercado	
Média	0.00049
Mediana	0.00040
Variância	0.0000017
Desvio Padrão	0.0004
Máximo	0.0020
Mínimo	0.00

Fonte: Elaborada pelo autor (2021)

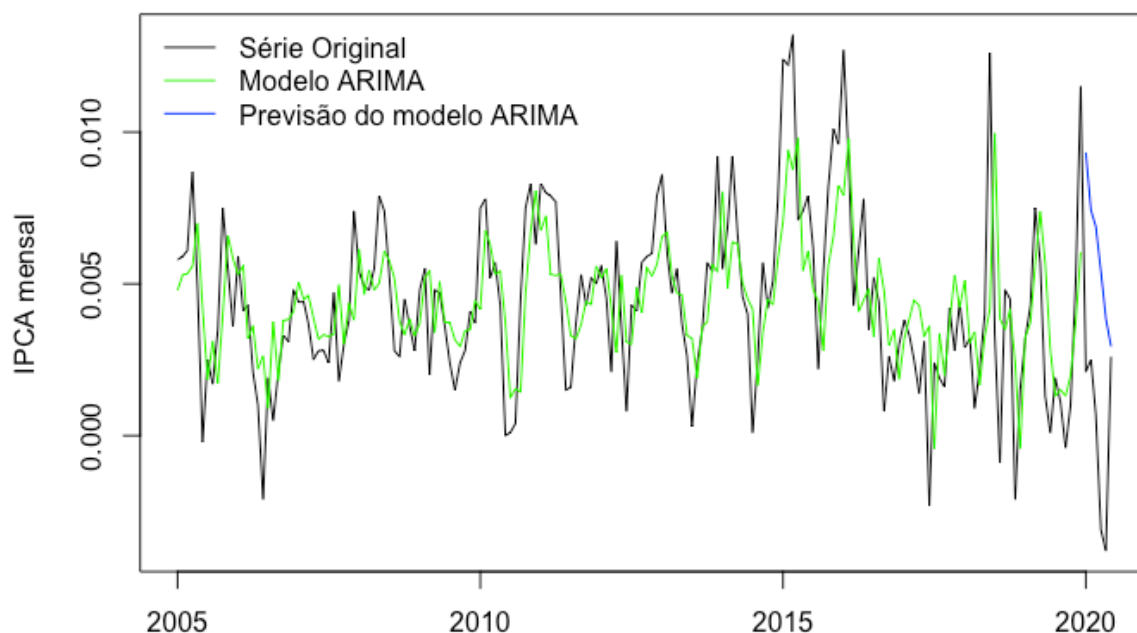
6.1.2 Simulação dos dados para períodos fora de análise

A segunda forma de análise proposta para o modelo ARMA(2,3) com média diferente de zero aplicado sobre a série IPCA mensal no período de janeiro de 2005 até dezembro de 2019 é a projeção do modelo para data fora do período de análise. Como a amostra total de dados eram de 180 meses, definiu-se que seria possível simular até seis meses a frente do período sem que os parâmetros iniciais deixassem de fazer sentido.

Assim, foram projetados através dos parâmetros definidos na Tabela 5.2 para o modelo ARIMA valores para os seis primeiros meses do ano de 2020, primeiros meses subsequentes ao fim do período inicial definido. O resultado, comparado também com os valores reais do IPCA para o mesmo período, estão apresentados na Figura 6.2.

Figura 6.2 - Simulação dos valores futuro para o modelo ARIMA da série IPCA mensal

Previsão de valores do IPCA mensal através do modelo ARIMA



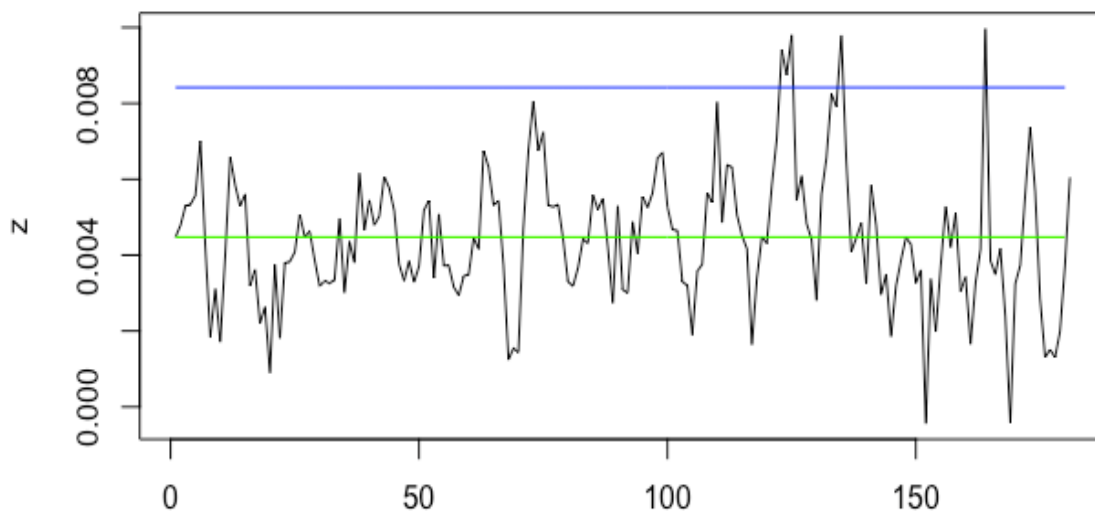
Fonte: Elaborado pelo autor (2021)

Percebe-se que, logo no início de 2020, houve um choque de inflação negativo, muito por conta do inesperado choque com a pandemia do Corona vírus. No período, dois dos seis dados (abril e maio de 2020) apresentaram valores abaixo de zero, algo que não aconteceu em nenhum outro momento na série inteira, inclusive com os valores negativos absolutos também os maiores desde o início da amostra, o que confirma o momento atípico do período estendido. Como esperado, o modelo não conseguiu replicar de forma confiável os valores reais e necessita do período de seis meses até aproximar-se novamente da inflação real e diminuir a diferença dos valores. Apesar disso, dada a grande dependência do modelo em relação ao primeiro valor exatamente anterior ao período desejado, o que mostra-se plausível para um modelo de inflação que tende a seguir uma tendência estabelecida para o período, o modelo ARIMA definido mostrou-se capaz de, mesmo quando colocado frente a período atípicos, adaptar-se para que os resultados continuem confiáveis, ao menos para um curto período de tempo até que os parâmetros tenham necessidade de serem redefinidos.

6.2 Gráfico de Controle

O resultado apresentado na Seção 5.5 está apresentado novamente na Figura 6.3, que mostra o resultado do gráfico de controle gerado após as simulações feitas para a série resultante da aplicação do modelo ARIMA sobre a série IPCA mensal.

Figura 6.3 – Gráfico de controle final para o modelo ARIMA da série IPCA mensal



Fonte: Elaborado pelo autor (2021)

A partir do gráfico de controle obtido, há duas principais análises a serem feitas: a primeira delas em relação aos parâmetros definidos para o gráfico, com o intuito de entender principalmente o que mostra o valor de lambda (λ) assumido para o gráfico; também é possível fazer uma observação da amostra de valores que foi detectado como fora de controle, que procura descobrir se houve na prática um descontrole inflacionário ou se houve uma falha na previsão do modelo ARIMA e consequentemente do gráfico de controle. As duas análises serão aprofundadas adiante.

6.2.1 Análise do parâmetro Lambda (λ)

Conforme as equações (4.18) e (4.19), o fator lambda define o peso que é dado aos valores passados na definição do fator z_t que irá ser aplicado no gráfico de controle, de forma que $\lambda(1 - \lambda)^i$ é a componente genérica do decaimento do peso que o fator tem ao longo do tempo i . Assim, quanto menor o fator, maior será o peso dado para os valores passados.

Conforme informações do capítulo 5.5 do presente trabalho de formatura, foi fixado para a simulação e definição dos limites do gráfico de controle um parâmetro lambda (λ) igual a um, que indica que não há pesos para valores passados nos gráficos, assim, z_t , portanto, é igual ao valor presente x_t . Isso leva a duas principais conclusões sobre o gráfico de controle:

- i) Como o fator x_t , definido através do modelo ARMA(2,3) com média diferente de zero já tem um peso muito relevante dos fatores imediatamente anteriores, através dos coeficientes de médias móveis e

auto regressivos, indicados na Tabela 5.2, não houve a necessidade do uso novamente dos dados passados durante a simulação dos limites de controle;

- ii) o gráfico de controle EWMA definido inicialmente encaixa-se nas definições do gráfico de controle de Shewhart após o resultado, na medida que realiza o julgamento do processo baseado somente na observação atual da amostra de dados. Em tese, o uso do gráfico de Shewhart, conforme Montgomery (2004), perde a sensibilidade a pequenas mudanças, justamente pela desconsideração da sequência passada de valores, fato que inclusive fez com que se define-se o uso de um gráfico de controle EWMA. Porém, é interessante notar que essa desvantagem intrínseca dos gráficos de Shewhart foi superada pelo fato do modelo ARIMA aplicado já conter as informações passadas.

6.2.2 Estudo dos pontos fora de controle

Conforme a Figura 6.3, há cinco valores que foram detectados pelo gráfico como fora de controle, todos apresentados na Tabela 6.3.

Tabela 6.3 – Pontos fora de controle

Indexador	Data	z_t	IPCA	$ \Delta $
122	02/2015	0,94	1,22	0,28
123	03/2015	0,88	1,32	0,44
124	04/2015	0,98	0,71	0,27
134	02/2016	0,98	0,90	0,08
163	07/2018	1,00	0,33	0,67

Fonte: Elaborada pelo autor (2021)

Conforme dados da tabela, a diferença entre os valores do IPCA real para os sinais do gráfico de controle, com a exceção do quarto ponto, são relativamente altas em relação ao que foi mostrado como média das mesmas diferenças na Tabela 6.1. Isso demonstra, em um primeiro momento, que os períodos tendem a estarem relacionados com choques inflacionários, em que a previsão do modelo é dificultada, mas que mostram por outro lado inflações fora de controle.

Em análise individual dos pontos, percebe-se que três deles são subsequentes, entre fevereiro e abril de 2015. De toda a amostra de dados, de fato o ano em questão foi aquele que

acumulou maior alta no índice de IPCA, também apresenta cinco das oito maiores altas mensais da amostra de 180 meses. Assim, é possível classificar o momento como um período de instabilidade para a inflação, dado que as variações estavam significativamente acima da média histórica.

Em relação ao segundo mês de 2016, percebe-se que a pressão inflacionária relativa ao ano anterior, analisada no parágrafo anterior, ainda surtia efeitos nos primeiros meses do ano seguinte, de forma que o segundo mês do ano, ponto identificado como fora de controle, apresentou a décima primeira maior alta mensal para toda a amostra de 180 dias. Assim, é possível identificar também que a variação mensal identificada pelo gráfico como fora de controle estava de fato acima do normal.

Por fim, o ponto relativo ao sétimo mês de 2018 está inserido em um momento em que também houve um descontrole inflacionário, porém, no mês exatamente anterior ao indicado pelo gráfico. No sexto mês de 2018 o índice apresentou o terceiro maior aumento mensal da amostra de 180 meses, mas que logo arrefeceu-se para os meses seguintes, indicou um choque de inflação isolado, diferente dos identificados anteriormente. Esta situação foi percebida pelo gráfico de controle com um lag de 1 mês, de forma que para os meses seis e sete de 2018 também foram identificadas as duas maiores diferenças entre o IPCA real e do modelo de toda a amostra de 180 dias.

6.3 Considerações Finais

Conforme observado ao longo das seções 6.1 e 6.2 o resultado da aplicação prática do uso de séries temporais e gráficos de controle para o monitoramento da inflação apresentou resultados que satisfatórios, de forma que o objetivo assumido para o presente trabalho de formatura está contemplado.

Em relação ao modelo ARIMA, o principal destaque do modelo é a capacidade de prever a inflação com uma precisão aceitável, visto que nem modelos usados na prática atualmente pelo mercado apresentam uma boa exatidão dos resultados. Acima de tudo, a capacidade do modelo ARIMA definido de adaptar-se a momentos de choques inflacionários inesperados, encaixa-se a nova realidade é um ponto forte do modelo e aumenta sua durabilidade e a confiança dos dados gerados mesmo em momentos de instabilidade.

Adiante, o gráfico de controle Shewhart mostrou-se capaz de identificar de forma quase 100% precisa pontos que remetem inflações de fato acima do esperado e cumpre seu objetivo principal.

7 CONCLUSÃO

O último capítulo do presente trabalho de formatura, além de trazer um breve resumo de todo o trabalho realizado, faz uma análise crítica do projeto e busca-se identificar como o resultado atinge o objetivo definido inicialmente. Por fim, são apresentadas sugestões de estudos futuros com base nas maiores dificuldades e limitações do caso prático desenvolvido em cima do índice de inflação brasileiro.

7.1 Desenvolvimento e resultados obtidos

O presente trabalho de formatura consiste no uso de conteúdos teóricos de séries temporais e de gráficos de controle para o monitoramento do índice oficial de inflação brasileiro, o IPCA.

Em um primeiro momento, além de resumir a parte teórica que envolve o trabalho no Capítulo 4, o trabalho de formatura também fez um resumo da importância e dos desafios que são encontrados no Brasil para monitorar a inflação, no Capítulo 2. Também foi apresentado no Capítulo 3 um breve histórico do crescimento e desenvolvimento de modelos quantitativos no âmbito do mercado financeiro, tendência ainda pouco explorada em território nacional quando comparada ao cenário externo.

O desenvolvimento do caso prático, feito no Capítulo 5, tem como objetivo estabelecido trabalhar a série dos retornos mensais do IPCA através de modelos ARIMA e replicar a série através de fatores auto regressivos e de médias móveis, capazes prever o comportamento da série. Busca-se também identificar através de limites de controle, sinais de instabilidade do mercado e consequentemente pontos de atenção para a tomada de decisões na prática.

No Capítulo 6 estão apresentadas análises em cima dos resultados obtidos, que mostram que as ferramentas definidas e utilizadas para monitoramento da inflação foram capazes de identificar os principais momentos de instabilidade do índice ao longo do período de estudo estabelecido.

7.2 Análise da adequação ao objetivo do trabalho

Conforme objetivos estabelecidos na Seção 1.4, “deseja-se obter um modelo que gere sinais confiáveis para um melhor controle do IPCA e que possa ser implantado no dia a dia da

gestora como forma de auxiliar na tomada de decisões ligadas à inflação, índice que é indexador de ao menos 85% dos ativos da carteira do Endurance Juros Reais”.

No Capítulo 6, apresentou-se a análise dos resultados obtidos do caso prático, que mostrou que o modelo ARMA(2,3) com média diferente de zero estabelecido para a previsão da inflação da série IPCA mensal entre o período de janeiro de 2005 a dezembro de 2019 é capaz de adaptar-se a momentos de instabilidade de mercado e identifica os picos de inflação, de forma a reestabelecer-se novamente na normalidade após o fim do período de alta do índice.

Ainda adiante, o controle estatístico do modelo através dos limites de controle estabelecidos para um gráfico de controle de Shewhart identificaram com extrema precisão os principais períodos de instabilidade da inflação do período. Os 5 pontos identificados como fora de controle estão de fatos entre os principais meses do período em que a inflação esteve descontroladamente alta.

Com base no objetivo estabelecido e na análise dos resultados obtidos, pode-se confirmar que as ferramentas de aplicação prática são capazes de trazer resultados numéricos consistentes. Mesmo que existam algumas limitações nos resultados, as ferramentas quantitativas mostram-se extremamente importantes capazes de auxiliar o monitoramento da inflação.

REFERENCIAS BIBLIOGRÁFICAS

ARESTIS, P. DE PAULA, L. F. FERRARI-FILHO, Fernando. **A nova política monetária: uma análise do regime de metas de inflação no Brasil**. Revista Economia e Sociedade, v. 18, n. 1, p. 1, 2009.

BCB. O que é inflação. **Política Monetária do Banco Central do Brasil**, 2021. Disponível em: <https://www.bcb.gov.br/controleinflacao/oqueinflacao>. Acesso em: 10 jun. 2021.

BCB. Índices de Preços. **Política Monetária do Banco Central do Brasil**, 2021. Disponível em: <https://www.bcb.gov.br/controleinflacao/indicepreco>. Acesso em: 10 jun. 2021.

BOX, G. E. P.; JENKINS, G. M. **Time Series Analysis: Forecasting and Control**. 1a ed. San Francisco: Holden-Day, 1970.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time Series Analysis: Forecasting and Control**. 3a ed. Englewood Cliffs: Prentice Hall, 1994.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time Series Analysis: Forecasting and Control**. 3a ed. Englewood Cliffs: Prentice Hall, 1994.

CASA CIVIL DA PRESIDÊNCIA DA REPÚBLICA. Lei Nº 12.431, de 24 de junho de 2011. **Subchefia para Assuntos Jurídicos da Casa Civil da Presidência da República**, 24 jun. 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112431.htm. Acesso em: 31 mai. 2021.

CECILIANO, V. **Avaliação de Desempenho de fundos quant.** Monografia de Final de Curso - Departamento de Ciências Econômicas da Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro, 2017.

CVM. Consulta Consolidada de Fundo. **Comissão de Valores Mobiliários**, 2019. Disponível em: https://cvmweb.cvm.gov.br/swb/default.asp?sg_sistema=fundosreg. Acesso em: 10 jun. 2021.

DICKEY, D. A.; FULLER, W. A. **Distribution of the Estimators for Autoregressive Time Series with a Unit Root**. Journal of the American Statistical Association, v. 74, n. 366, p. 427, 1979.

HOLLAND, M. MORI, R. **Dinâmica da Inflação no Brasil e os Efeitos Globais**. Brasília, DF. Associação dos Centros e Pós-Graduação em Economia – ANPEC, v. 11, n. 3, p. 649, 2010.

INFO MONEY. 5 Investimentos Arelados à Inflação para proteger o seu portfólio da alta do IPCA. **Info Money**, 9 jun. 2021. Disponível em: <https://www.infomoney.com.br/onde-investir/5-investimentos-atrelados-a-inflacao-para-protger-o-seu-portfolio-da-alta-do-ipca/>. Acesso em: 31 ago. 2021.

INFO MONEY. Fundos Quantitativos: conheça os gestores que ganham com a volatilidade do mercado. **Info Money**, 18 mai. 2019. Disponível em:

<https://www.infomoney.com.br/onde-investir/fundos-quantitativos-conheca-os-gestores-que-ganham-com-a-volatilidade-do-mercado/>. Acesso em: 31 ago. 2021.

JOURNEY CAPITAL. Endurance Juros Reais. Fundo de Inflação com investimento em projetos de infraestrutura. **Journey Capital Adm de Recursos Ltda**, 2019. Disponível em:

<https://www.journeycapital.com.br/endurance-juros-reais/>. Acesso em: 10 jun. 2021.

KUMIEGA, A.; VAN VLIET, B. E. **Automated finance: The assumptions and behavioral aspects of algorithmic trading**. Journal of Behavioral Finance, v. 13, n. 1, p. 51–55, 2012.

MONTGOMERY, D. C. **Introdução ao Controle Estatístico da Qualidade**. 4a ed. Rio de Janeiro: Editora LTC, 2004.

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

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. 2a ed. São Paulo: Editora Blucher, 2006.

PAG SEGURO. Fundos de debêntures incentivadas: saiba como investir e entenda como funcionam. **Pag Seguro Pag Bank**, 13 abr. 2021. Disponível em:

<https://blog.pagseguro.uol.com.br/fundos-de-debentures-incentivadas-saiba-como-investir-e-entenda-como-funcionam/#rmcl>. Acesso em: 31 mai. 2021.

PASTORE, A. C. PINOTTI, M. C. **Inflação e Estabilização: Algumas Lições da Experiência Brasileira**. Revista Brasileira de Economia, v. 53, n.1, p. 3, 1999.

SACHSIDA, A. **Inflação, desemprego e choques cambiais: uma revisão da literatura sobre a Curva de Phillips no Brasil**. Revista Brasileira de Economia, v. 67, n. 4, p. 549, 2013.

VALOR INVESTE. Cresce a busca por títulos atrelados à inflação no Tesouro Direto. É bom comprar?. **Valor Investe**, São Paulo, 21 jun. 2021. Disponível em:

<https://valorinveste.globo.com/produtos/renda-fixa/tesouro-direto/noticia/2021/06/21/cresce-a-busca-por-titulos-atrelados-a-inflacao-no-tesouro-direto-e-bom-comprar.ghtml>. Acesso em: 31 ago. 2021.

XP INVESTIMENTOS. Por que você deveria investir um pedaço do seu patrimônio em fundos de debêntures incentivadas?. **Análise e Seleção de Fundos XP**, 31 mar. 2021.

Disponível em: <https://conteudos.xpi.com.br/fundos-de-investimento/relatorios/porque-voce-deveria-investir-um-pedaco-do-seu-patrimonio-em-fundos-de-debentures-incentivadas/>.

Acesso em: 31 mai. 2021.

THE WALL STREET JOURNAL. Behind the Market Swoon: The Herdlike Behavior of Computerized Trading. **The Wall Street Journal**, 25 dez. 2018. Disponível em: <https://www.wsj.com/articles/behind-the-market-swoon-the-herdlike-behavior-of-computerized-trading-11545785641>. Acesso em: 31 ago. 2021.

ZUCKERMAN, G. **The Man Who Solved the Market**. How Jim Simons launched the Quant Revolution. 1a ed. Penguin Random House LLC, 2019.

APÊNDICE A: CÓDIGO EM R PARA APLICAÇÃO DO CASO PRÁTICO

```
#Importando as bibliotecas necessárias
```

```
library('forecast')
```

```
library('tseries')
```

```
library('FitAR')
```

```
library('readxl')
```

```
library("sarima")
```

```
library("astsa")
```

```
#Leitura e ajuste da base de dados IPCA
```

```
base_ipca <- read_excel(file.choose())
```

```
colnames(base_ipca) <- c("data", "fechamento", "retorno")
```

```
base_ipca$fechamento <- as.numeric(base_ipca$fechamento)
```

```
base_ipca$retorno <- as.numeric(base_ipca$retorno)
```

```
base_ipca$data <- as.Date(base_ipca$data)
```

```
base_ipca_completa <- base_ipca[1:(nrow(base_ipca)-12),]
```

```
base_ipca <- base_ipca[1:180,]
```

```
#Grafico da série e definição da série temporal
```

```
plot(base_ipca$data, base_ipca$retorno, type = "l", main = "Série IPCA mensal", ylab = "IPCA mensal", xlab = "")
```

```
serie_ipca = ts(base_ipca[, c('retorno')])
```

```
#autocorrelações e autocorrelações parciais
```

```
Acf(serie_ipca, main = 'ACF para a série IPCA mensal')
```

```
Pacf(serie_ipca, main = 'PACF para a série IPCA mensal')
```

```
#testes de estacionariedade e normalidade
```

```

adf.test(serie_ipca, alternative = "stationary")
hist(serie_ipca, xlim = c(-0.01,0.03), breaks = 25, ylab = "Frequência",
     main = "Histograma da série IPCA mensal")
qqnorm(serie_ipca, main = "Q-Q Plot da curva normal para a série IPCA mensal",
      ylab = "Quantis da amostra", xlab = "Quantis teóricos")
qqline(serie_ipca)
shapiro.test(serie_ipca)

```

```

#auto simulação para encontrar o modelo ARIMA
ipca = auto.arima(serie_ipca, stepwise = F)
print(ipca)

```

```

#cálculo do modelo e estimação dos parâmetros
ipca = arima(serie_ipca, order = c(2,0,3))
print(ipca)
abs(ipca$coef)/sqrt(diag(ipca$var.coef))

```

```

#testes dos resíduos do modelo ARIMA
tsdisplay(residuals(ipca), lag.max = 15,
      main = "Resíduos do modelo ARIMA para a série IPCA mensal")
qqnorm(resid(ipca),
      main = "Q-Q Plot da curva normal para os resíduos do modelo ARIMA",
      ylab = "Quantis da amostra", xlab = "Quantis teóricos")
qqline(resid(ipca))
shapiro.test(resid(ipca))
cx <- Box.test(resid(ipca), type = "Ljung",lag=30,fitdf=7)
LBQPlot(resid(ipca), lag.max = 15)

```

```

#grafico do modelo vs serie original
plot(base_ipca$data,base_ipca$retorno,type = "l",

```

```

    main = "Comparação das séries original e gerada através do modelo ARIMA",
    xlab = "", ylab = "IPCA mensal")
lines(base_ipca$data,fitted(ipca),col = "green")
legend("topleft", legend = c("Série Original","Modelo ARIMA"), lwd = 1, bty = "n", col =
c("black", "green"))

```

```
#analizando os resultados
```

```

base_ipca$retorno_modelo <- fitted(ipca)
base_ipca$diferenca <- abs(base_ipca$retorno - base_ipca$retorno_modelo)
mean(abs(base_ipca$diferenca))
median(abs(base_ipca$diferenca))
sd(base_ipca$diferenca)
as.numeric(var(base_ipca$diferenca))
max(base_ipca$diferenca)
min(base_ipca$diferenca)

```

```
#projetando valores dos primeiros 6 meses de 2020
```

```

inflacao_modelo = predict(ipca, n.ahead = 6)
ipca_serie_completa = ts(base_ipca_completa$retorno,
    start = c(1995,1), frequency = 12)
plot(base_ipca_completa$data,base_ipca_completa$retorno, type = "l",
    ylab = "IPCA mensal", xlab = "",
    main = "Previsão de valores do IPCA mensal através do modelo ARIMA")
lines(head(base_ipca_completa$data, 180), fitted(ipca), col = "green")
lines(tail(base_ipca_completa$data, 6), inflacao_modelo$pred, col="blue")
legend("topleft", legend = c("Série Original","Modelo ARIMA","Previsão do modelo
    ARIMA"), lwd = 1, bty = "n", col = c("black", "green", "blue"))

```

```
#coeficientes do modelo para poder fazer a simulação
```

```

ar1 = ipca$coef[["ar1"]]
ar2 = ipca$coef[["ar2"]]

```

```

ma1 = ipca$coef[["ma1"]]
ma2 = ipca$coef[["ma2"]]
ma3 = ipca$coef[["ma3"]]
intercept = ipca$coef[["intercept"]]
sd = sd(residuals(ipca))

```

```

#parametros iniciais da simulacao
ARLinini = 18 #1.5ano
N = 1000
lam = 1
L = 0.001
step = 0.00001
stop = F
arls = c()

```

```

#simulacao para limite de controle
while (stop == F) {
  soma_arl = 0
  print(L)

```

```

  for (i in c(1:N)){
    x <- sarima.sim(ar = c(ar1,ar2), ma = c(ma1,ma2,ma3), sd = sd) + intercept
    z0 = mean(x)
    z = c(z0)
    mi = mean(x)
    t = 0
    fora = F

```

```

    while (fora == F) {
      t = t + 1
      zt = lam*x[t] + (1-lam)*z[t]
      z = append(z,zt)

```

```

#limites de controle para cada observacao
LSC <- L

#observação está fora dos limites
if (zt>LSC) {
  soma_arl = soma_arl + t
  fora = T}
}
}

#ARL médio do L utilizado
arl_sim = soma_arl / N
arls = append(arls,arl_sim)

#ARL médio for maior que o definido, a simulação termina
if (arl_sim >= ARLini){
  stop = T}
else{
  L = L + step}

print(arl_sim)
}

#parametros para testar o ARL
k = 1
step = 0.05
arls = c()

#simulacao para teste do ARL
while (k <= 5) {
  soma_arl = 0

```

```

print(k)
for (i in c(1:N)){

  x <- sarima.sim(ar = c(ar1,ar2), ma = c(ma1,ma2,ma3), sd=k*sd) + intercept
  z0 = mean(x)
  z = c(z0)

  t = 0
  fora = F
  while (fora == F){
    t = t + 1
    zt = lam*x[t] + (1-lam)*z[t]
    z = append(z,zt)

    #limites de controle para cada observacao
    LSC <- L

    #observacao esta fora dos limites
    if (zt>LSC) {
      soma_arl = soma_arl + t
      fora = T
    }
  }

  #calculos do ARL
  arl_sim = soma_arl / N
  arls = append(arls,arl_sim)
  k = k + step
  print(arl_sim)
}

#ARL x k, para verificar se esta dentro do esperado
plot(arls, type="l")

```

```

z0 = intercept
z = c(z0)
x = fitted(ipca)

#limites de controle para cada observacao
lscs = c()
lcs = c()

#pontos fora dos limites de controle
maior = 0

for (i in c(1:length(fitted(ipca)))){
  zt = lam*x[i] + (1-lam)*z[i]
  z = append(z,zt)

  #calculo dos limites de controle
  LSC <- L
  lscs = append(lscs,LSC)
  LC = intercept
  lcs = append(lcs,LC)

  #verificacao se o ponto estar dentro ou fora dos limites
  if (zt > LSC) {maior = maior + 1}
}

#grafico da serie juntamente com os limites de controle
plot(z, x=c(1:181),type="l")
lines(lscs,type="l",col="blue")
lines(lcs,type="l",col="green")

arl_serie = length(fitted(ipca)) / (maior)
print(arl_serie)

```