

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS  
Instituto de Ciências Econômicas e Gerenciais

**Enzo Antunes Souza**

**A PANDEMIA E O SETOR DE VIAGENS AÉREAS NO BRASIL: Uma análise  
econométrica**

Belo Horizonte

2024  
**Enzo Antunes Souza**

**A PANDEMIA E O SETOR DE VIAGENS AÉREAS NO BRASIL: Uma análise  
econométrica**

Trabalho de Conclusão apresentado na modalidade Monografia ao Curso de Ciências econômicas do Campus Coração Eucarístico como requisito parcial para obtenção do título de bacharel em Ciências econômicas.

Orientador: Prof. Duval Magalhães

## RESUMO

O presente projeto de monografia tem como ideia central analisar observar a série histórica de demanda por vôos aéreos ao longo do tempo e traçar um comparativo entre o momento pré e pós-pandemia através da estimação de um modelo ARIMA e a observação da diferença dos modelos em ambos os períodos de tempo, utilizando dados disponibilizados pela ANAC (Agência Nacional de Aviação Civil). Além disso, também será feita a análise dos dados, visando entender as mudanças e padrões que podem ser observados. **(Quando tiver os resultados finais, irei refazer conforme solicitado na última entrega)**

Palavras-chave: Pandemia, mercado de aviação civil, séries temporais, modelos econométricos, análise de dados.

**LISTA DE TABELAS**

Tabela 1 – Métricas do transporte aéreo.....	14
Tabela 2 – Padrões teóricos de AFC e PAFC nos modelos.....	28

## LISTA DE ABREVIATURAS E SIGLAS

ARMA	Autoregressive-moving-average (Modelo auto-regressivo de média móvel)
ARIMA	Autoregressive integrated moving average (Modelo auto-regressivo de média móvel)
COVID-19	Coronavirus disease 2019 (Doença por coronavírus 2019)
SARS	Severe Acute Respiratory Syndrome (Síndrome respiratória aguda)
TCC	Trabalho de Conclusão de Curso
ANAC	Agência Nacional de Aviação Civil
ABEAR	Associação Brasileira das Empresas Aéreas
AFC	Função de autocorrelação
PAFC	Função de autocorrelação parcial

## SUMÁRIO

1	INTRODUÇÃO.....	6
1.1	Aviação civil no Brasil e o impacto da COVID .....	6
2	REFERENCIAL TEÓRICO .....	11
2.1	Métricas da aviação civil no Brasil.....	11
2.2	Séries temporais.....	13
2.3	Modelos de séries temporais.....	19
2.4	Metodologia Box-Jenkins .....	21
3	METODOLOGIA.....	25
3	REFERÊNCIAS .....	35

# 1 INTRODUÇÃO

## 1.1 Aviação civil no Brasil e o impacto da COVID

O marco legal da aviação civil (lei Nº 7.565, de 19 de dezembro de 1986) foi feito a partir das normas e práticas recomendadas pela Organização da Aviação Civil Internacional, agência da Organização das Nações Unidas responsável por promover o desenvolvimento seguro e ordenado da aviação civil mundial, sendo formada por 193 Estados Membros. De acordo com o marco, o mercado de aviação civil possui um conceito amplo, que abrange um conjunto de atividades, tendo como principal segmento o transporte aéreo, que se subdivide em passageiros, carga e mala postal, além de também poder ser subdivido em regular (linhas permanentes) e não regular (possui caráter esporádico, como vôos charter e táxi aéreo). Também de acordo com o marco legal da aviação civil, o Código Brasileiro de Aeronáutica segmenta a aviação comercial doméstica em três categorias: transporte aéreo regular, transporte aéreo não regular e serviços especializados, sendo que a primeira depende de concessão, enquanto as demais dependem de autorização<sup>1</sup>. Por fim, temos a distinção entre vôos domésticos e vôos internacionais, sendo os domésticos aqueles onde o local de partida e chegada ocorrem no mesmo país e os internacionais, como o nome diz, ocorrem em países diferentes. O foco dessa monografia, será apenas no transporte aéreo regular de passageiros, tanto domésticos, quanto internacionais. No ano de 2023, essa modalidade transportou mais de 113 milhões de passageiros, de acordo com o Anuário de Transporte Aéreo da ANAC (2024), o que mostra a importância da mesma para a economia e descentralização das atividades no país.

Como foi descrito por Rosa Bergiante (2013), mesmo com aparente expansão em número de empresas e vôos, o setor globalmente vem enfrentando dificuldades no século XXI, devido à sua dependência a diversas externalidades e à pequena margem de lucro. Como exemplo dessas externalidades, Rosa cita os ataques que atingiram Nova York e Washington em 11 de Setembro de 2001, a crise de 2008 e,

---

<sup>1</sup> De acordo com Russar (2010) a concessão é um contrato administrativo onde o Poder Público delega a execução de serviços públicos a terceiros. Já a autorização, é um ato onde o Poder Público delega a um particular a exploração de serviço público de maneira precária, ou seja, essa autorização pode ser revogada a qualquer momento sem a necessidade de indenização prévia. A concessão é oferecida a aeroportos e companhias aéreas para que possam oferecer o serviço de transporte aéreo, já a autorização se enquadra mais em casos como taxi aéreo por exemplo.

curiosamente, questões sanitárias devido aos efeitos da SARS em 2003 na China. Mais recentemente, o setor volta a passar por grande dificuldade relacionada a questões sanitárias, porém agora com a pandemia do vírus COVID-19.

Ao analisarmos o mercado brasileiro atual, vemos que o setor é altamente concentrado, muito pela natureza do negócio, mas também pelas dificuldades enfrentadas ao longo do tempo pelas empresas. De acordo com o Painel de Indicadores do Transporte Aéreo de 2022 feito pela ANAC (2023), no mercado doméstico, 3 empresas representam mais de 99% da demanda por vôos domésticos, porém, apenas cerca de 23% no mercado internacional.

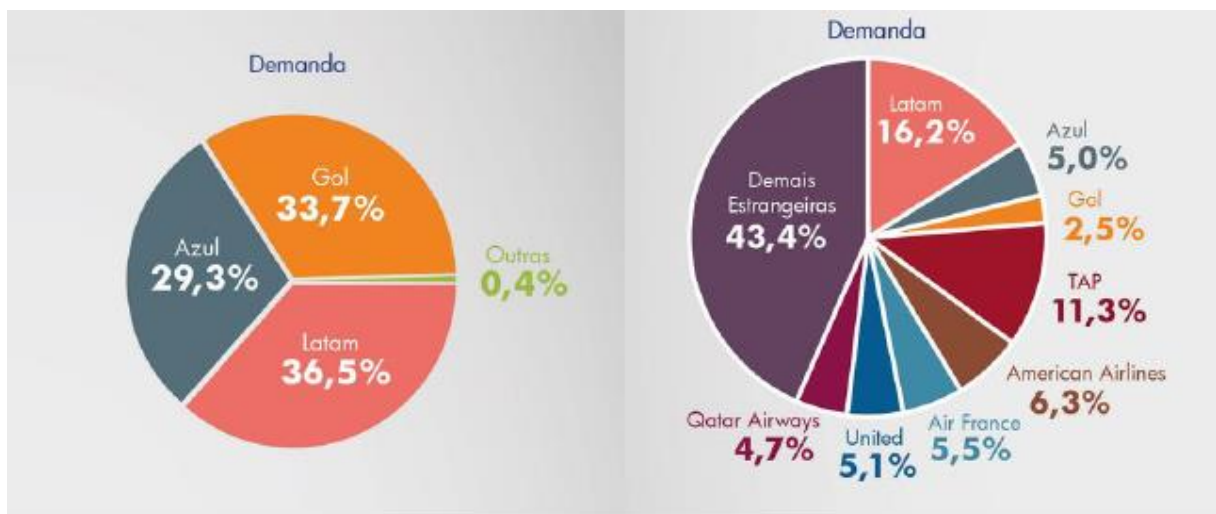


Figura 1 – Demanda por vôos domésticos vs vôos internacionais. Fonte: ANAC.

Além disso, podemos ver a quantidade de custos envolvidos nos vôos, que vão de combustíveis até seguros e tarifas, fazendo com que existam diversas externalidades que podem modificar a margem de lucro das companhias.



Figura 2 – Custos e despesas dos Serviços Aéreos em 2022. Fonte: ANAC



Apesar disso, na última década o setor vinha conseguindo se consolidar, apresentando um crescimento acelerado da demanda, impulsionada tanto pelo crescimento econômico do país, quanto por quedas nas tarifas médias, como aponta Leonardo Ribeiro Pinguelli (2021) em seu estudo sobre os impactos da pandemia no setor aéreo. De acordo com o autor, entre 2004 e 2011, o volume de passageiros cresceu à uma taxa de 11,8% ao ano, porém, a partir da década de 2010 o cenário econômico do país se modificou e, depois de oito anos em crescimento acelerado, o setor se depara com um crescimento de apenas 1% entre 2012 e 2013, como podemos observar nos gráficos abaixo:

**Figura 13 - Volume de Passageiros Transportados no Brasil (milhões)**

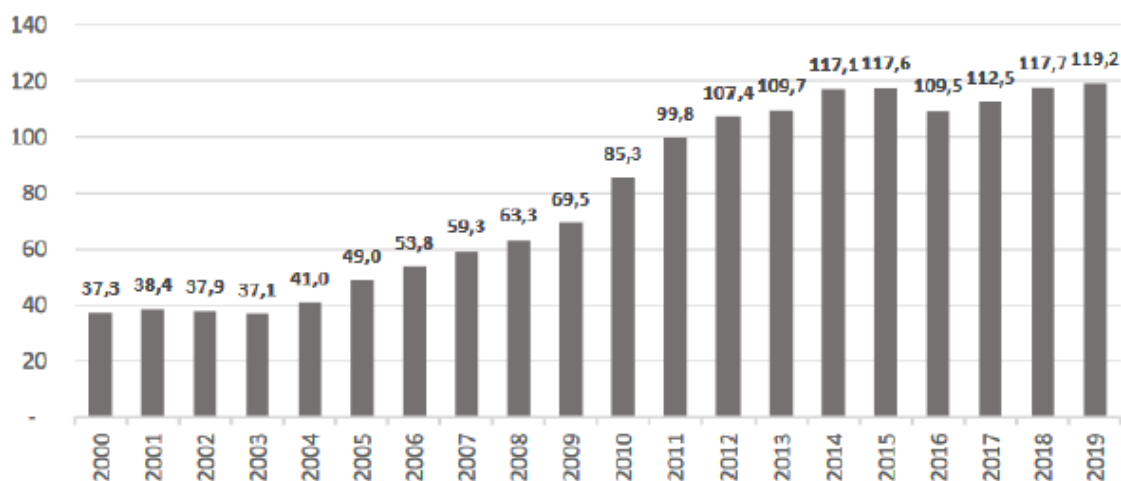


Figura 3 - Fonte: Pinguelli, Leonardo (2021) – A pandemia da COVID-19 e as companhias aéreas brasileiras

**Figura 7 - Evolução da demanda doméstica (RPK), PIB e tarifa média no Brasil**

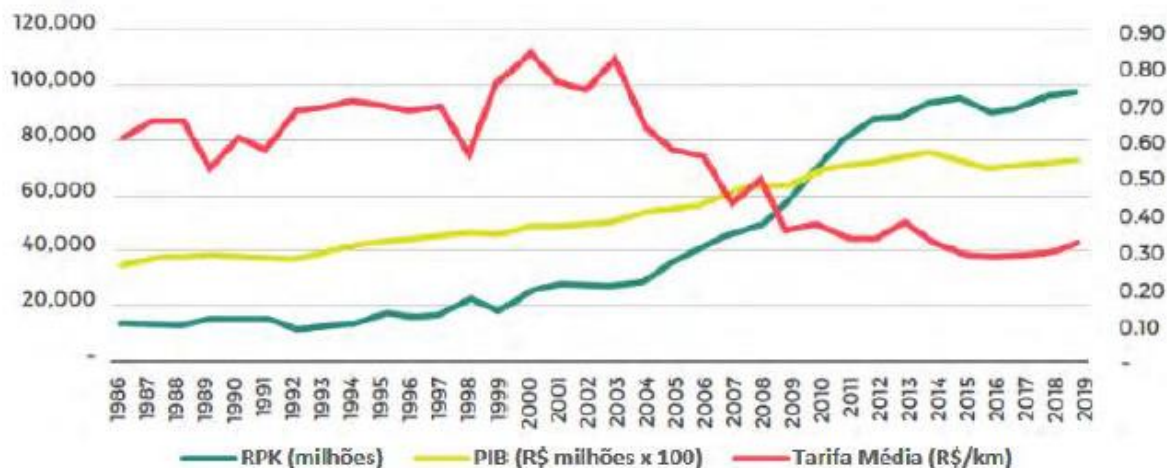


Figura 4 - Fonte: ABEAR

Sobre os impactos da pandemia no setor, é possível observar o quão lesado foi o mesmo. De acordo com o relatório Redes e Fluxo do Território – Ligações aéreas 2019-2020, o transporte de passageiros teve uma redução de 53% de 2019 para 2020 e até hoje o setor não se recuperou completamente, já que, de acordo com a ANAC, ao compararmos dados de 2019 (pré-pandemia) com a de 2022, vemos que a variação na demanda foi de -7,3% para o mercado doméstico de -29,7% para o mercado 10 internacional, fazendo com que fosse gerado só nesse último ano um prejuízo de R\$2,2 bilhões. A ABEAR, em seu último relatório sobre o panorama geral do setor, classifica os anos de 2020 e 2021 como os dois piores momentos da história da aviação brasileira. Para se ter uma ideia do tamanho do problema, em 2021, a demanda doméstica foi equivalente ao verificado em 2010 e, pior ainda no setor internacional, sendo equivalente ao verificado em 1992.

#### EVOLUÇÃO DOS PASSAGEIROS TRANSPORTADOS PAGOS EM VOOS DOMÉSTICOS E INTERNACIONAIS NO BRASIL (REAL X ESTIMADO)



Figura 5 - Fonte: ABEAR

A pandemia e lockdown causado pelo covid-19 foram extremamente prejudiciais para diversos setores em todo mundo, mas, principalmente para o setor de turismo e viagens aéreas, setor esse que, mesmo com todas as dificuldades enfrentadas em 2019, ainda assim foi responsável por 1,4% do PIB brasileiro, 60 mil empregados diretos e 1,4 milhões de outros postos de trabalhos em atividades correlacionadas indiretamente de acordo com relatório do ONTL (Observatório Nacional de Transporte e Logística de 2023), mostrando a relevância do mesmo para a sociedade como um todo.

Dada tamanha importância, se tornam relevantes estudos para entender sua demanda, além da velocidade de recuperação no mundo pós-pandemia, uma vez que o setor enfrentou uma queda de 53% no número de passageiros em 2020 de acordo com estudo do IBGE. Complementando, é possível ir mais além e criar e comparações com as séries pré-pandemia, visando entender se os padrões de consumo são mantidos e se sim, na mesma proporção, buscando investigar em que passo está a recuperação desse importante setor da nossa economia.

A monografia possui como objetivo geral entender, através da análise de dados e econometria, qual a real situação do mercado aéreo brasileiro pós pandemia, buscando definir como se dá a sua recuperação e entender os impactos da pandemia no mesmo.

Já como objetivos específicos, a monografia busca analisar a sazonalidade das séries temporais de números de vôos aéreos comerciais pré e pós-pandemia no Brasil; Investigar numericamente qual foi o impacto da pandemia no número de vôos aéreos no Brasil; Analisar as séries temporais atuais de números de vôos aéreos pós-pandemia (2022-2024), bem como entender sua sazonalidade e criar paralelos com as séries pré-pandemia (2015-2018), estimar um modelo ARIMA com os dados pré-pandemia e comparar suas previsões com os valores reais que vemos em cada mês de 2024.

Por fim, a pesquisa sugerida possui justificativas tanto teóricas quanto práticas. No que diz respeito aos motivos teóricos, podemos dizer que a aplicação de modelos econométricos de séries temporais é altamente relacionado com o problema que a pesquisa busca resolver, ampliando o acervo de exemplos onde a econometria pode ser aplicada e mostrando o poder do ferramental estatístico e tecnológico (já que a manipulação de dados e criação do modelo será feita utilizando a linguagem de programação Python3) para a solução de problemas econômicos. Através da linguagem Python3, será criado um modelo ARIMA (modelo auto-regressivo de médias móveis) utilizando a metodologia Box&Jenkins e, com ele, será feita a previsão dos valores futuros do número de passageiros para cada mês de 2024 e a comparação da previsão do modelo com os valores reais.

Em relação aos motivos práticos, foi visto sobre a importância do setor para a sociedade brasileira e o quanto ele foi prejudicado pela pandemia, tornando necessária uma análise de como o mercado se encontra atualmente, possibilitando um debate de como o governo pode ajudar o mesmo a se recuperar, caso necessário.

Além disso, existe também um motivo pessoal para a criação da monografia, já que trabalho com modelos estatísticos e programação a 2 anos e meio, e escolhi o tema justamente por poder juntar todos os conhecimentos que adquiri na minha jornada profissional aos meus conhecimentos em economia adquiridos ao longo da graduação.

Portanto, a pesquisa pode trazer diversas contribuições para o debate econômico e político brasileiro, além de ser um caso excelente de junção de diversas ciências (programação, economia e estatística) para entender um fenômeno econômico.

## **2 REFERENCIAL TEÓRICO**

Para iniciarmos a discussão sobre os impactos da COVID-19 no setor aéreo e a análise quantitativa em si, é de suma importância a introdução de alguns conceitos relacionados ao modelo de negócio e à análise estatística de séries temporais para facilitar o entendimento do que foi feito.

### **2.1 Métricas da aviação civil no Brasil**

Nos dados disponibilizados pela ANAC, existem diversas métricas para reunir informações e facilitar a interpretação dos dados, além de termos específicos referentes ao negócio. Nessa seção, iremos apresentar uma tabela com os conceitos e selecionar quais métricas serão utilizadas ao longo de nossa análise.

Inicialmente, é importante separar os dois tipos de etapas do voo: a etapa básica (flight stage) e a etapa combinada (On flight origin and destination – OFOD). De acordo com a ANAC “As etapas básicas são aquelas realizadas pela aeronave desde a sua decolagem até o próximo pouso, independente de onde tenha sido realizado o embarque ou o desembarque do objeto de transporte, apresentando a movimentação de cargas e passageiros entre os aeródromos de origem e destino da aeronave. É a operação de uma aeronave entre uma decolagem e o próximo pouso, ou seja, é a ligação direta entre dois aeródromos”, já as etapas combinadas “identificam as partes de aeródromos de origem, onde houve o embarque do objeto de transporte, e destino, onde houve o desembarque do objeto de transporte, independente da existência de aeródromos intermediários, atendidos por determinado

voo. É a etapa de voo vista com foco no objeto de transporte (pessoas e/ou cargas), com base no embarque e desembarque nos aeródromos relacionados”. Portanto, podemos observar que na etapa OFOD, os dados são registrados de acordo com a origem e destino do voo, independente das suas escalas. Por fim, aeródromo pode ser considerada qualquer superfície que possua infraestrutura destinada à aterragem, decolagem e movimentação de aeronaves.

Isso posto, segue abaixo a tabela com os termos e seus significados:

*Tabela 1 - KPIs de análise de dados de vôos aéreos*

Termo	Definição
Payload	É a capacidade total de peso na aeronave, expressa em quilogramas, disponível para efetuar o transporte de passageiros, carga e correio.
ASK (Avaible seat kilometer)	Refere-se ao volume de Assentos Quilômetros Oferecidos, ou seja, a soma do produto entre o número de assentos oferecido e a distância das etapas.
RPK (Revenue passenger kilometer)	Refere-se ao volume de Tonelada Quilômetro Oferecida, ou seja, a soma do produto entre o payload, que é a capacidade total de peso disponível na aeronave, expressa em quilogramas, disponível para efetuar o transporte de passageiros, carga e correio, e a distância das etapas, dividido por 1.000.
ATK (Avaible tonne kilometer)	Refere-se ao volume de Tonelada Quilômetro Oferecida, ou seja, a soma do produto entre o <i>payload</i> , que é a capacidade total de peso disponível na aeronave, expressa em quilogramas, disponível para efetuar o transporte de passageiros, carga e correio, e a distância das etapas, dividido por 1.000.
RTK (Revenue tonee kilometer)	Refere-se ao volume de Toneladas Quilômetros Transportadas, ou seja, a soma do produto entre a distância das etapas e os objetos pagos transportados expressos em quilogramas (carga, correio, passageiro e bagagem) dividido por 1.000. Cada passageiro possui o peso estimado de 75 kg para empresas brasileiras. No caso das empresas estrangeiras, que não enviam dados de bagagem à ANAC, o peso total de passageiros e bagagens é estimado multiplicando-se a quantidade total de passageiros por 90;

## 2.2 Séries temporais

Uma série temporal é um conjunto de observações dos valores que uma variável assume em diferentes momentos do tempo. Esses dados podem ser coletados a intervalos regulares, como diariamente, semanalmente, mensalmente, etc. (GUJARATI, 2011) e são de extrema importância para a análise de fenômenos que se repetem ao longo do tempo. A Figura 5 deste documento é um exemplo de série temporal, onde, no caso, está sendo observada a variável número de passageiros transportados. Uma série temporal pode ser classificada quanto ao número de elementos na mesma.

Se o conjunto de observações no tempo for finito, então ela é considerada discreta, porém, se o conjunto gerado for infinito ou não enumerável, então ela é contínua. Uma série contínua é denotada por:  $\{X(t): t \in T\}$ ,  $T=\{t:t_1 < t < t_2\}$ , já no caso discreto  $\{X_t: t \in T\}$ ,  $T=\{t_1, \dots, t_n\}$ . (EHLERS, 2007).

Outra análise que pode ser feita em relação às séries temporais é quanto a sua tendência, podendo ser crescente, decrescente, etc. No caso abaixo, vemos a série temporal crescente que mostra o número de casos de COVID-19 no Brasil por semana epidemiológica:



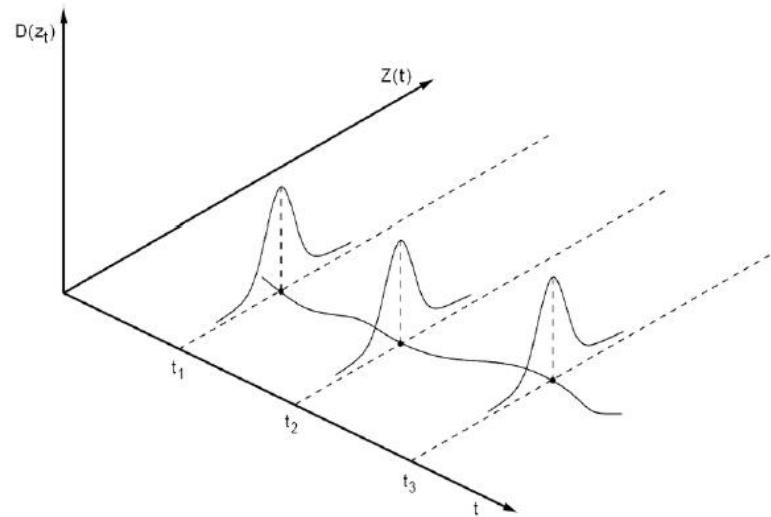
Figura 6 - Fonte: Ministério da saúde

### 2.3.1 Processos estocásticos

Os modelos utilizados para descrever séries temporais são processos estocásticos, que são processos cujos valores se dão através de uma distribuição de probabilidade. Podemos definir um processo estocástico da seguinte forma:

- Seja  $T$  um conjunto arbitrário. Um processo estocástico é uma família  $Z=\{Z(t), t \in T\}$ , tal que, para cada  $t \in T$ ,  $Z(t)$  é uma variável aleatória. (TOLOI, 2006)

Dessa forma, um processo estocástico é uma família de variáveis aleatórias, onde o conjunto  $T$  é tomado como o conjunto dos números inteiros ou do conjunto dos reais e, para cada  $t \in T$ ,  $Z(t)$  será uma variável real. A imagem abaixo ilustra um processo estocástico interpretado como uma família de variáveis aleatórias:



*Figura 7- Processo estocástico como uma família de variáveis aleatórias*

Na figura, para cada período de tempo  $t$ , temos uma variável aleatória com sua distribuição de probabilidade, sendo possível que essa distribuição seja diferente ao longo dos instantes  $t$ , porém o ideal é que ela seja a mesma para todo  $t \in T$ . No estudo de séries temporais, elas são tratadas como processos estocásticos, onde cada valor assumido pela variável de análise ao longo do tempo é determinado por uma variável aleatória (TOLOI, 2006).

### 2.3.2 Estacionariedade

Um dos pressupostos comumente utilizado em análises de séries temporais é de que a série analisada seja estacionária, ou seja, sua média e variância não variam sistematicamente ao longo do tempo (GUJARATI, 2011), ou seja, ela avança aleatoriamente em torno de uma média constante. O processo estacionário pode ser estritamente estacionário ou fracamente estacionário.

- Séries estritamente estacionária

Um processo é estritamente estacionário se todos os momentos de sua distribuição de probabilidade, e não apenas os dois primeiros (ou seja, a média e a variância), são

invariantes ao longo do tempo (GUJARATI, 2011). Isso ocorre quando a distribuição de probabilidade conjunta de  $X(t_1), \dots, X(t_k)$  é a mesma de  $X(t_1+n), \dots, X(t_k+n)$ , ou seja, o deslocamento da origem dos tempos por uma quantidade  $n$  não tem efeito na distribuição conjunta que portanto depende apenas dos intervalos entre  $t_1, \dots, t_k$  (BARBOSA, 2012). Na prática, é comum definir a estacionariedade de uma forma menos restrita, já que é difícil utilizar a definição de estacionariedade estrita (EHLERS, 2007).

- Série fracamente estacionária

Um processo estocástico diz-se fracamente estacionário ou estacionário de segunda ordem quando sua função média é constante e sua função de autocovariância depende apenas da defasagem:

- $E\{Z(t)\} = \mu(t) = \mu$ , constante para todo  $t \in T$ ;
- $\gamma(t_1, t_2) = COV\{Z(t_1), Z(t_2)\}$  é uma função de  $|t_1 - t_2|$ ; (TOLOI, 2006)

- Série não estacionária

Por fim, existem diversos tipos de não-estacionariedade, porém, no nosso estudo focaremos nos processos não-estacionários homogêneos, que são aqueles cujo nível e/ou inclinação mudam com o decorrer do tempo e, mesmo que as séries não possuem um valor fixo para sua média, seu comportamento em diferentes períodos de tempo é similar, como podemos ver na figura abaixo:

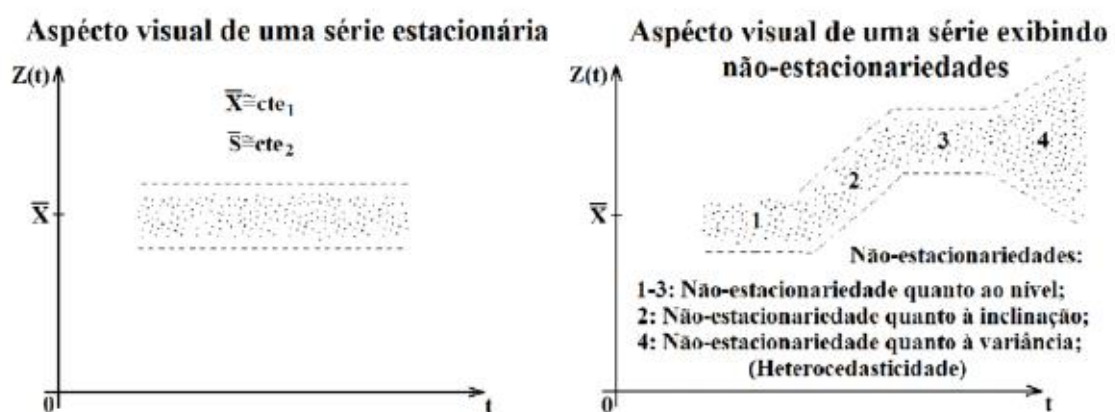


Figura 8 - Série estacionária vs não estacionária



Os mesmos podem se tornar estacionários através de sucessivas diferenciações (TOLUI, 2006). Em geral, as séries com comportamento inconstante são mais comuns, especialmente no setor de economia.

### 2.3.3 Processos estocásticos integrados

Como dito anteriormente, em geral, as séries mais comuns no setor da economia são não-estacionárias, portanto, necessitamos torná-las estacionárias, já que a maior parte dos procedimentos de análise de séries temporais pressupõe que as mesmas são estacionárias. Visando esse objetivo, as diferenciações ou integrações são técnicas que consistem em diferenciar séries estacionárias até que as mesmas se tornem estacionárias. Se uma série tem de ser  $d$  vezes diferenciada para torná-la estacionária, dizemos que ela é uma série temporal integrada de ordem  $d$  e denotamos da seguinte forma:  $Yd \sim Id$ . A maior parte das séries temporais econômicas são  $I(1)$ , ou seja, necessitam apenas uma diferenciação para se tornarem estacionárias (GUJARATI, 2011).

Podemos definir a primeira diferença dessa nossa série  $Y_t$  por:

$$\Delta Y_t = Y_t - Y_{t-1},$$

a segunda por

$$\Delta^2 Y_t = \Delta[Y_t - Y_{t-1}] \Delta^2 Y_t = Y_t - 2Y_{t-1} + Y_{t-2},$$

e assim por diante.

### 2.3.4 Testes de estacionariedade

Como dito anteriormente, para utilizarmos a maior parte das ferramentas econométricas de análise de séries temporais, devemos ter em mãos uma série temporal estacionária. Para isso, é importante saber se uma série é ou não estacionária. Essa dúvida é retirada com os testes de estacionariedade, que nesse estudo serão utilizados o teste de correlograma e a análise gráfica.

- Análise gráfica

Pela análise gráfica, é possível ter uma pista inicial sobre a natureza provável da série. Para exemplificar, vamos analisar a série abaixo:

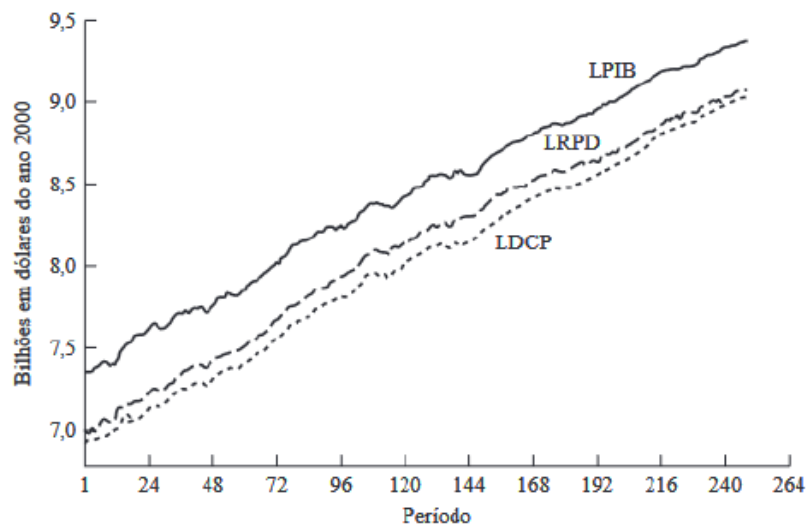


Figura 9 - Logaritmos de PIB, RPD e DCP reais, EUA (1947-2007) em bilhões de dólares

Nela, vemos que, ao longo dos trimestres, o logaritmo do PIB tem aumentado, indicando uma tendência de aumento e que sugere que talvez a média do PIB tenha se alterado nesse período. (GUJARATI, 2011). A análise gráfica é importante, mas serve apenas para uma ideia inicial do comportamento da série e não exclui a necessidade de execução de testes formais, como a análise da função de correlação (FAC) e correlograma, que serão explicados a seguir.

- Função de correlação e correlograma

Um teste de estacionariedade é baseado na função de correlação (FAC), onde uma FAC com defasagem  $k$ , denotada por  $\rho_k$  é definida como:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{covariância com defasagem } k}{\text{variância}}$$

em que a covariância (ou autocovariância) com defasagem  $k$  é a covariância entre os valores  $Y_t$  e  $T_{t+k}$  e a variância é uma medida de dispersão que mostra o quão distante o valor está da média. (GUJARATI, 2011) Logo, dessa divisão obtemos um valor  $\rho_k$  que varia de -1 a 1, como qualquer coeficiente de correlação faz e, ao colocarmos em uma gráfico versus  $k$ , temos o correlograma.

Nesse caso, vamos precisar calcular tanto a covariância da amostra da nossa série temporal com defasagem  $k$  ( $\hat{\gamma}_k$ ), quanto a variância da amostra ( $\hat{\gamma}_0$ ). Ambas denotadas por:

$$\hat{\gamma}_k = [\sum (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})] / n, e$$

$$\hat{\gamma}_0 = [\sum(Y_t - \bar{Y})^2]/n$$

onde  $n$  é o tamanho da amostra e  $\bar{Y}$  é a média da amostra. Com isso, temos que a função de correlação amostral é  $\hat{\rho}_k = \hat{\gamma}_k / \hat{\gamma}_0$ , que é a razão das duas grandezas demonstradas acima. É importante ressaltar que a correlação parcial mede a correlação entre as observações que estão separadas por  $k$  períodos, ignorando as correlações nas defasagens intermediárias. Já a função de correlação, leva em consideração todas as defasagens de  $t_1$  a  $t_k$ .

Ao obter esses valores, é possível obter o correlograma amostral, como o exemplificado a seguir:

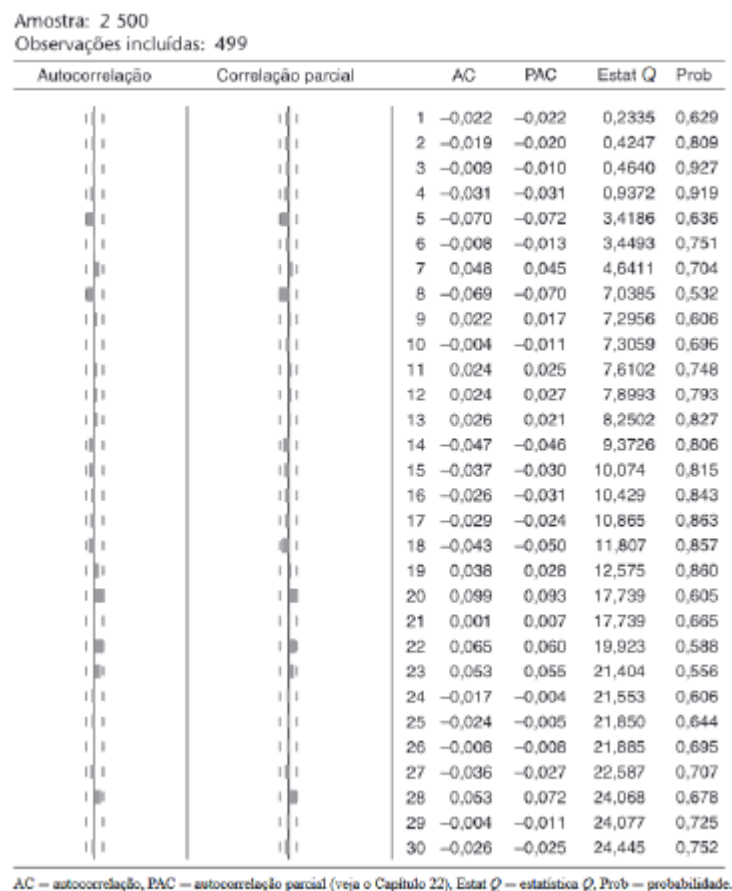


Figura 10 - Correlograma de termo de erro  $u$  com 30 defasagens

Nesse diagrama, na coluna “Autocorrelação” e “Correlação parcial”, a linha vertical representa o eixo zero e as observações à direita desse eixo são positivas e à esquerda são negativas. É possível perceber nesse diagrama que sua autocorrelação gira em torno de zero em várias defasagens, exemplificando um correlograma de uma série temporal estacionária, então podemos dizer que a série temporal que originou esse correlograma provavelmente é estacionária. Já nesse

outro caso, vemos um correlograma característico de uma série temporal não estacionária, onde o valor do coeficiente de autocorrelação cai muito lentamente ao decorrer das defasagens.

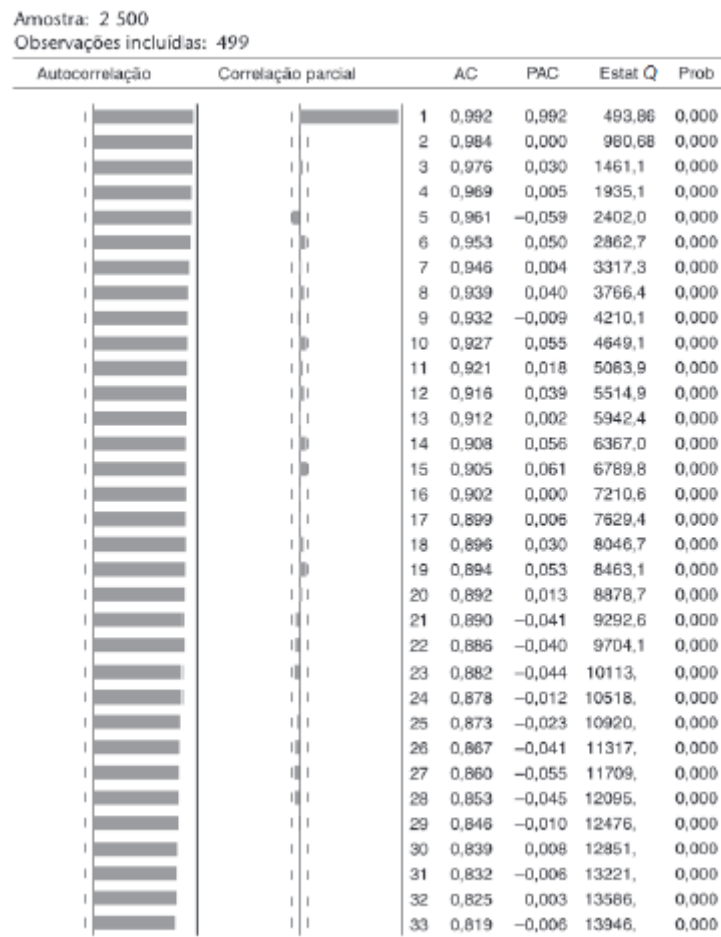


Figura 11 - Correlograma de uma série temporal de passeio aleatório

De acordo com Gujarati e Porter, para escolhermos a extensão da defasagem ao criarmos o correlograma, devemos seguir uma regra empírica de que a escolha do número de defasagens deve estar entre um terço até um quarto da extensão da série temporal. Por fim, podemos perceber que é possível descobrir se uma série é estacionária ou não utilizando tanto a análise gráfica quanto o correlograma. A partir do momento que determinamos a estacionariedade da série, podemos prosseguir com nossa análise.

## 2.3 Modelos de séries temporais

Os modelos para séries temporais podem ser classificados segundo o número de parâmetros envolvidos, podendo ser de dois tipos: modelos paramétricos, para os quais o número de parâmetros é finito, ou modelos não paramétricos, que envolvem um número infinito de parâmetros. Dentro dos modelos paramétricos, os mais frequentemente utilizados são os modelos auto-regressivos integrados e de médias

móveis (ARIMA), de onde podemos obter os modelos auto-regressivos e de médias móveis (ARMA) como sendo um caso particular de modelo ARIMA.

### 2.3.1 Modelos autoregressivos (AR)

Nesse tipo de modelo, o valor de  $Y$  no período  $t$  depende do seu valor no período anterior e de um termo aleatório. Os valores de  $Y$  são expressos como desvios com base em um valor médio. Logo, esse modelo afirma que o valor atual de  $Y$  são alguma proporção da média mais perturbações aleatórias no período  $t$  (GUJARATI, 2011). Cada modelo AR tem uma ordem de acordo com o número de defasagens utilizadas para defini-lo. Por exemplo, o modelo:

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_2(Y_{t-2} - \delta) + u_t,$$

onde  $\delta$  é a média de  $Y$  e  $u_t$  é um erro aleatório não correlacionado com média zero e variância constante (chamamos esse tipo de erro de ruído branco), podemos dizer que  $Y_t$  segue um processo autorregressivo estocástico de segunda ordem AR(2), já que apresenta duas defasagens (o valor atual de  $Y$  depende apenas do seu valor nos dois período prévios).

### 2.3.2 Modelos de média móvel (MA)

O processo AR descrito acima também pode nos ajudar a desvendar os modelos de média móvel, que pode ser descrito por uma combinação linear de termos de erro. Por exemplo, podemos ter o seguinte modelo AM:

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1},$$

onde  $\mu$  é uma constante e  $u_t$ , como anteriormente, é um termo de erro estocástico de ruído branco. Nesse caso, temos que  $Y$  segue um processo de média móvel de primeira ordem, já que utilizamos o erro de apenas um período no passado.

### 2.3.3 Modelos autoregressivos de média móvel (ARMA)

Continuando, é possível utilizar os dois modelos descritos anteriormente, já que uma série pode ter características de ambos os processos. Desses casos, surge o ARMA. A série a seguir pode ser denotada como um ARMA(1,1):

$$Y_t = \theta + \alpha_1 Y_{t-1} + \beta_0 u_t + \beta_1 u_{t-1},$$

já que apresenta tanto um termo autoregressivo, quanto um termo de média móvel. Em geral, um processo ARMA( $p, q$ ) haverá termos autorregressivos  $p$  e termos de média móvel  $q$  (GUJARATI, 2011).

### 2.3.3 Modelos autoregressivos integrados de média móvel (ARIMA)

Os modelos anteriormente apresentados são baseados na suposição de que as séries temporais são estacionárias. Porém, como já foi discutido anteriormente, a maior parte das séries econômicas e de fenômenos sociais não são estacionárias, o que se cria a necessidade de incluir a integração no modelo para obtermos séries estacionárias (foi visto que uma série temporal não estacionária  $I(d)$  depois de  $d$  diferenciações se torna uma série estacionária). Nesse sentido, se é necessário a diferenciação  $d$  vezes para, após isso, ser aplicado um modelo ARMA( $p, q$ ), é dito que a série é ARIMA( $p, d, q$ ), ou seja, uma série temporal autoregressiva integrada de médias móveis, onde  $p$  é o número de termos autorregressivos,  $d$  o número de diferenciações que a série sofreu até se tornar estacionária e  $q$  o número de termos de média móvel (GUJARATI, 2011).

## 2.4 Metodologia Box-Jenkins

Uma metodologia bastante utilizada na análise de modelos paramétricos (número finito de parâmetros) é a metodologia Box-Jenkins, que nos ajuda a determinar os valores dos coeficientes  $p, d, q$  e, conseqüentemente, escolher qual tipo de modelo utilizar (AR, MA, ARMA ou ARIMA) e consiste em um passo a passo para ajustar esses modelos a um conjunto de dados. A metodologia consiste em quatro etapas iterativas que devem ser executadas na ordem em que se segue:

1. Identificação: Com a utilização do correlograma, descobrimos quais os valores apropriados de  $p, d, q$ .
2. Estimação: Nessa etapa é feita a estimação dos parâmetros dos termos autorregressivos e dos termos de média móvel incluídos no modelo. Pode ser feito pelo método dos mínimos quadrados simples, ou por vezes, por métodos de estimação não linear.
3. Verificação do diagnóstico: Após a escolha do modelo e estimação de seu parâmetro, é feita a avaliação do modelo para descobrir se ele é um modelo que realmente descreve bem a relação entre os dados. Nessa etapa, é possível utilizar diversos testes estatísticos para avaliar o modelo, porém, um dos testes mais simples é verificar se os resíduos estimados com base nesse modelo são ruídos brancos. Se sim, o ajuste pode ser aceito, porém, se não, é necessário recomeçar do primeiro passo. (GUJARATI, 2011)
4. Previsão: Em diversos casos, um modelo ARIMA é estimado visando prever valores do futuro, já que ele é extremamente consistente nessas previsões. Portanto, como parte do teste do modelo estimado, fazer previsões é interessante.

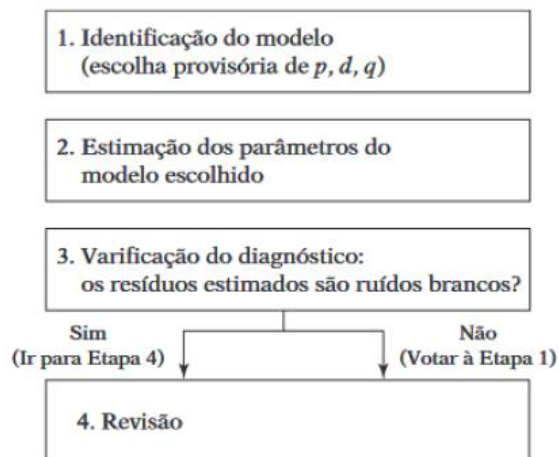


Figura 12 - O método Box-Jenkins. Fonte: GUJARATI, 2011

#### 2.4.1 Identificação do tipo de modelo

Para identificar um modelo, como dito anteriormente, é de grande ajuda a sua função de correlação amostral (ACF) e função de correlação parcial (PACF), uma vez

que as mesmas geram os correlogramas que serão analisados para determinar os parâmetros  $p, d, q$  do modelo. Partindo para um exemplo, é possível analisar a ACF e PACF abaixo:

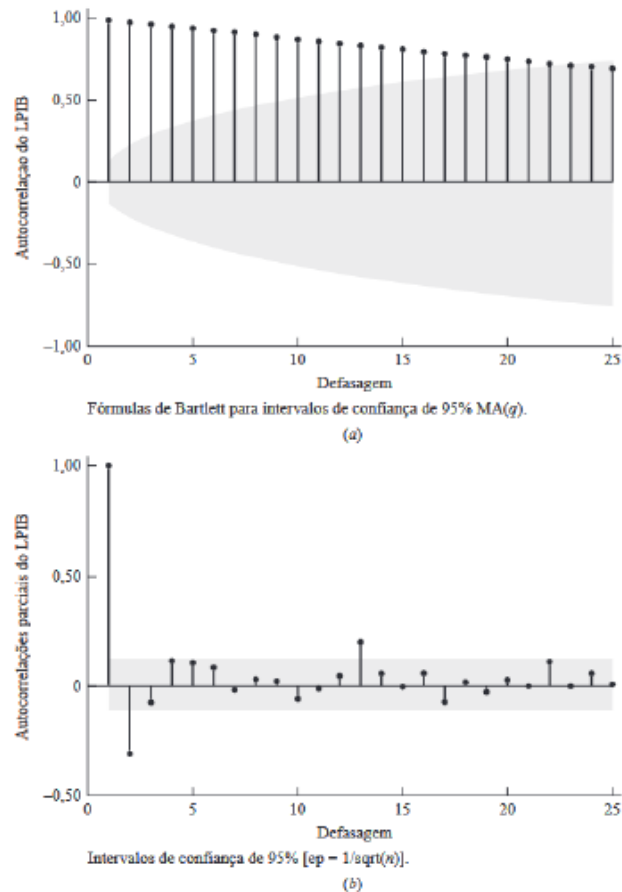


Figura 13 - ACF e PACF de exemplo. Fonte: GUJARATI, 2011

Nelas, é possível perceber que a ACF (imagem a) decai muito lentamente, chegando ao intervalo de confiança de 95% (mancha cinza) apenas a partir de 21 defasagens, logo, até a 21ª defasagem, os valores são estatisticamente diferentes de zero, o que, como visto anteriormente, indica que a série possivelmente não é estacionária. Nesse sentido, é necessário torná-la estacionária com diferenciações e, após a primeira diferenciação, é possível observar as seguintes ACF e PACF:



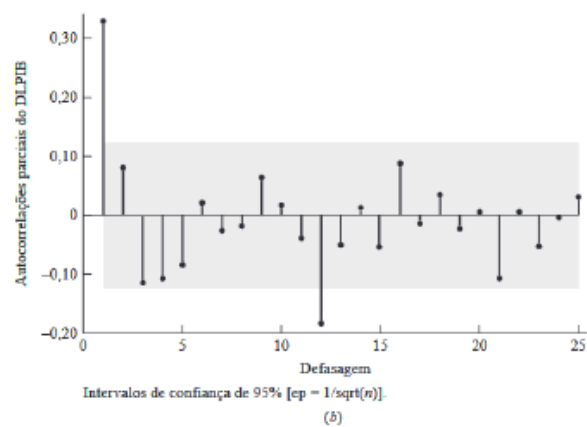
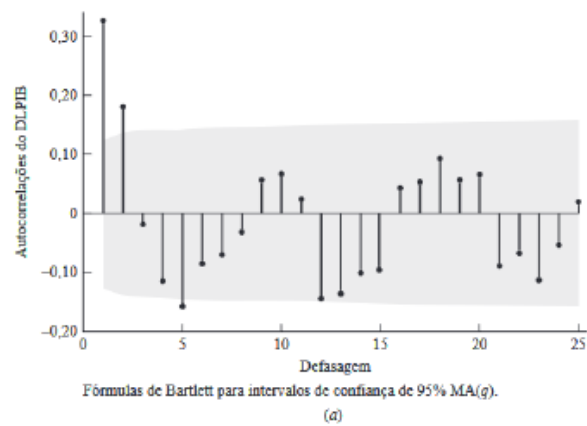


Figura 14 - ACF e PACF de exemplo após a primeira diferenciação

Após a primeira diferenciação, é possível perceber que a ACF segue um padrão de uma série estacionária, ou seja, rapidamente seu valor cai e, a partir da segunda defasagem apresenta valor estatisticamente igual a zero. Portanto, com a análise dos correlogramas, é possível identificar o primeiro parâmetro  $d$  do modelo, que, no caso, seria igual a 1.

Uma maneira comum de descobrir os outros dois parâmetros  $p$  e  $q$ , é observar os padrões que os modelos mais básicos, por exemplo AR(1), AR(2), MA(1), MA(2), ARMA(1), ARMA(2), seguem nos seus correlogramas da função de correlação e função de correlação parcial. Com isso, é importante analisar as referências para tirar as conclusões do valor dos parâmetros em cada caso.

No exemplo abaixo, é possível ver os padrões mais comuns para certos modelos:

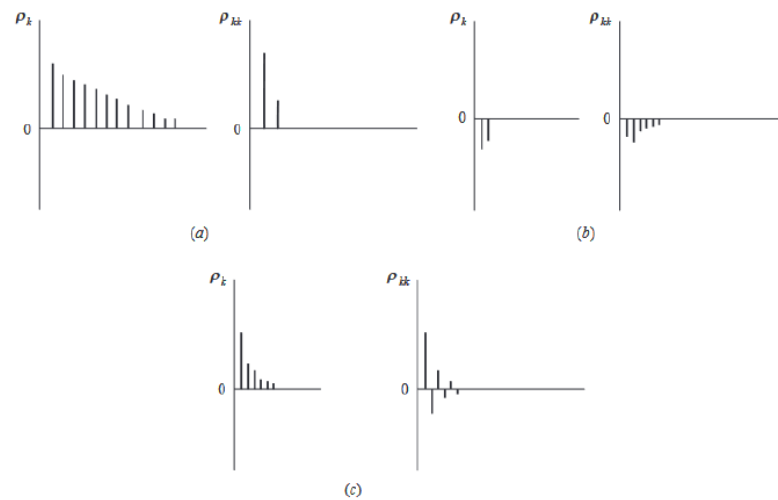


Figura 15 - ACF e PACF de alguns processos mais comuns: (a) AR(2); (b) MA(2); (c) ARMA(1,1). Fonte GUJARATI, 2011

Por fim, os padrões teóricos das ACF e PACF de cada tipo de modelo, como dito anteriormente, pode ajudar na escolha dos parâmetros através da comparação do padrão com a situação enfrentada com os dados selecionados:

Tabela 2 - Padrões teóricos de ACF e PACF

Tipo de modelo	Padrão típico de ACF	Padrão típico de PACF
AR(q)	Declina exponencialmente ou com padrão de onda senóide amortecida ou ambos.	Picos significativos até p defasagens.
MA(q)	Picos significativos até q defasagens.	Declina exponencialmente.
ARMA(p,q)	Queda exponencial	Queda exponencial.

### 3 METODOLOGIA

A análise será feita em cima dos dados disponibilizados pela ANAC e a métrica escolhida como variável dependente do modelo é o RPK, uma vez que essa métrica é a mais utilizada nas análises de demandas feitas tanto em outros trabalhos científicos, quanto em relatórios da própria ANAC. Como dito anteriormente, essa métrica é o volume de Passageiros Quilômetros Transportados (Revenue Passenger Kilometer).

A base de dados escolhida foi o Relatório de Demanda e Oferta, que é uma base mensal desde 2000, porém, para os objetivos desse projeto, serão utilizadas as datas de 2018-2019 para a série pré covid-19 e de janeiro de 2023 a março de 2024 para a série pós covid-19. A coluna escolhida para gerar a série histórica foi a `nr_rpk` (que indica o RPK) e o tipo de etapa utilizada na análise será a etapa básica (flight stage), que são aquelas que representam a operação de uma aeronave entre uma decolagem e o próximo pouso, independentemente se for o ponto final do passageiro ou não. Esse tipo de etapa foi selecionado, uma vez que o RPK é calculado em cada etapa básica.

Após conseguir os dados, é importante procurar por qualquer tipo de falta de informação ou dados faltantes e tratar isso da maneira que fizer mais sentido, uma vez que isso pode influenciar negativamente em nossos resultados, e separar as bases de dados entre as duas datas de referência desejadas (pré-covid e pós-covid) e selecionar as colunas do nosso interesse. Toda a manipulação de dados será feita com a linguagem de programação python e com a biblioteca pandas, referência mundial em manipulação de dados em formatos tabulares.

Com o dado em mãos e tratados, será feita uma análise gráfica e estatística em cima desses dados, buscando diferenças e semelhanças ao longo das duas séries temporais e buscando entender mais profundamente qual o impacto da pandemia e se o impacto ainda está presente ou se dissipou ao longo do tempo. A partir disso, será possível prosseguir para a metodologia Box-Jenkins para estimar um modelo de séries temporais na série pré-covid.

Inicialmente, na etapa de identificação do modelo, serão desenhadas as funções de correlação parcial e função de correlação e, a partir delas, serão observados os padrões teóricos apresentados para, finalmente, definir quais serão os parâmetros e tipo do modelo a ser estimado. Na etapa de estimação, será utilizado o método de máxima verossimilhança com a biblioteca `sci-kit learn` do python (biblioteca que oferece um conjunto de funções novas à linguagem), que nos permite criar diversos modelos, avaliá-los, criarmos previsões, etc. Com isso, conseguiremos obter nosso modelo estimado e, como próximo passo, será feita a verificação dos parâmetros do modelo para que seja garantido a significância estatística dos parâmetros. Para isso, serão observados os erros dos modelos para se determinar se são parecidos com ruídos branco e, além disso, serão utilizados testes formais para se determinar a significância.

Por fim, será feita uma previsão utilizando esse modelo, caso ele seja estatisticamente significativo, para os meses de janeiro de 2023 a março de 2024 (série pós-covid) e serão comparados os valores da previsão com os valores reais de RPK registrados pela ANAC.

## 4 APLICAÇÃO DA METODOLOGIA BOX JENKINS

Essa sessão será destinada à explicação do passo a passo que foi tomado para coletar e analisar os dados, além de estimar o modelo com seus parâmetros seguindo a metodologia Box-Jenkins. Todas as análises e gráficos foram geradas através da linguagem de programação *python3* e o código executado para chegar nos resultados estará disponível na plataforma *github*.

### 4.1 Análise exploratória e preparação dos dados

Inicialmente, foram coletados os dados no site da ANAC, selecionando voos de 01/01/2000 a 01/04/2024. Como a base possuía informações de cada voo, inclusive voos não regulares, foi necessário filtrar para termos apenas voos regulares e agrupar essas informações para termos o número de passageiros pagos mensalmente de 01/01/2000 a 01/04/2024. Com isso, chegamos a uma base de dados com as colunas “ANO”, indicando o ano de referência, “MÊS”, indicando o mês de referência e “PASSAGEIROS\_PAGOS”, indicando o número de passageiros que pagaram para voar naquele mês e ano. Após isso, buscando entender melhor o comportamento dessa série temporal pré-pandemia (foi selecionado de 2012 a 2019), foi feito o gráfico abaixo:

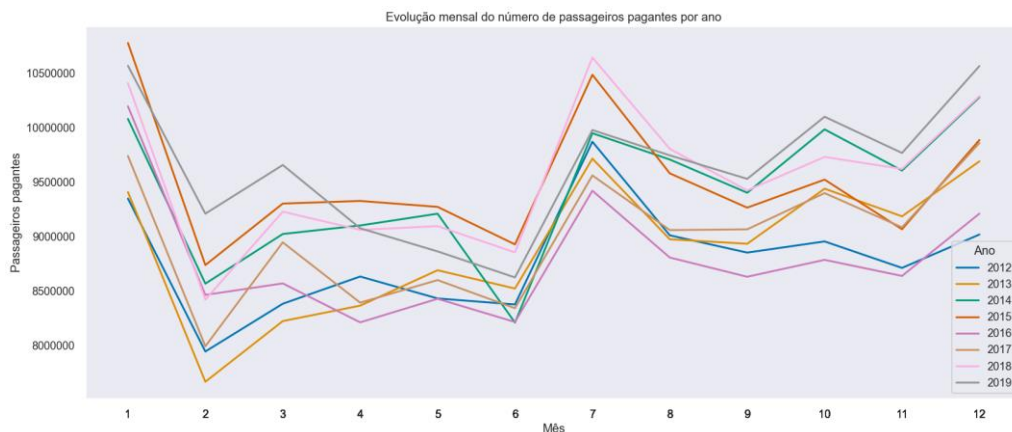


Figura 16 - Evolução mensal do número de passageiros pagantes por ano (2012 a 2019)

Nele, é possível perceber uma clara sazonalidade, uma vez que todos os anos mostram uma queda no mês de Fevereiro (fim das férias de verão) e aumento tanto no mês de Julho, quanto no mês de Dezembro (início das férias tanto do meio do ano quanto de verão).

O mesmo foi feito com a série pós-covid (foi selecionado 2020 a 2024):

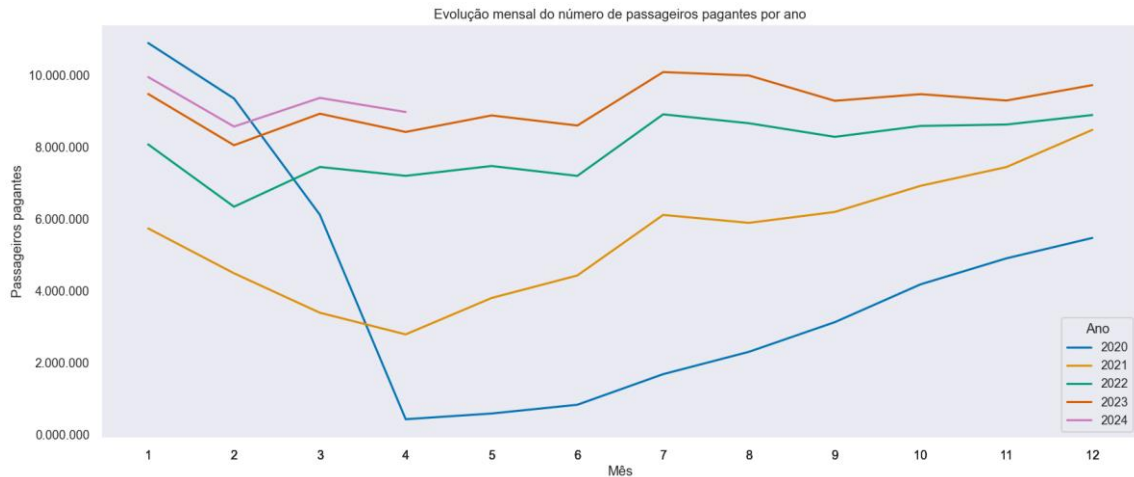


Figura 17- Evolução mensal do número de passageiros pagantes por ano (2020 a 2024)

Nela é possível claramente observar a queda no ano de 2020 relacionada ao *lockdown*. É possível perceber que nos anos seguintes o número de passageiros pagantes foi aumentando e, nos gráficos de 2022 a 2024 é possível ver os mesmos padrões da série pré-covid.

Outra forma de se observar a sazonalidade é através do gráfico abaixo:

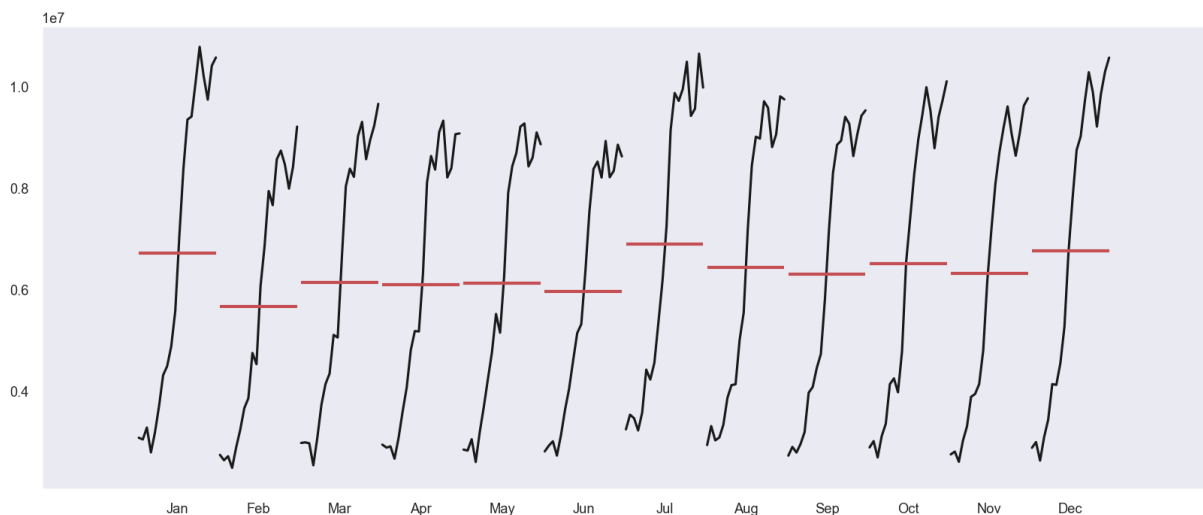


Figura 18 - Gráfico de sazonalidade da série total até o covid (2000 a 2019)

Nele, conseguimos perceber como a média (linha vermelha) varia de acordo com o mês de referência, tendo uma variação maior nos meses de Janeiro, Fevereiro, Julho e Dezembro. Ao se constatar que a série de fato apresenta sazonalidade, foi

necessário dessazonalizar a mesma para então aplicarmos a metodologia *Box-Jenkins*. Para isso, utilizei a decomposição multiplicativa da série, onde o componente sazonal é considerado constante de um ano para o outro e que usa como base a fórmula  $Y_t = S_t \times T_t \times E_t$ , onde  $Y_t$  é o dado observado no período  $t$ ,  $S_t$  é a componente sazonal em  $t$ ,  $T_t$  é a componente tendência em  $t$  e  $R_t$  é o erro em  $t$  (GIARETTA, 2017). Esse método foi escolhido, uma vez que é o mais apropriado em casos onde a variação da sazonalidade aumenta ao decorrer da série, que é o que observamos na Figura 16 (quanto maior o ano, maior é a variação que a sazonalidade causa). Ao dessazonalizar, foi obtida a seguinte série, onde é possível se observar uma série laranja mais suave e com menos aumentos e quedas bruscas em relação à série sazonalizada em azul:



Figura 19 - Série de número de passageiros pagantes antes e após ser dessazonalizada

## 4.2 Identificação

Essa etapa possui como objetivo identificar quais ordens de  $p$ ,  $d$  e  $q$ . Lembrando que  $p$  é o termo autoregressivo,  $q$  o termo de médias móveis e  $d$  o termo de diferenciações.

Primeiramente, foi verificado a estacionariedade da série através do teste de raiz unitária ADF (Augmented Dickey-Fuller) através da função *adfuller* da biblioteca *statsmodels*, chegando aos valores de -2,6 para a estatística de Dickey-Fuller e um p-valor de 0,093. Dado que o valor crítico a um nível de significância de 0,05 é de -2,874, a hipótese nula (há uma raiz unitária ou a série temporal é não estacionária) não é rejeitada, já que ela só é rejeitada se  $DF < z$ , provando que a série temporal não é estacionária.

Portanto, para prosseguir, foi necessário executar uma diferenciação e refazer o teste ADF na nova série diferenciada. A diferenciação foi feita com a função *diff* da biblioteca *statsmodels*. No novo teste, foi encontrado uma estatística Dickey-Fuller de -2,74 e um p-valor de 0,07. Dado que o valor crítico a um nível de significância de 0,05 é de -2,874, a hipótese nula ainda não foi rejeitada (já que  $DF > z$ ), sendo necessária mais uma diferenciação. Após a segunda diferenciação, foi observado um valor de -6,916 para a estatística Dickey-Fuller, podendo, finalmente, rejeitar a hipótese nula do teste e provar que a nova série gerada é estacionária com média zero. Com isso, o valor  $d$  do modelo ARIMA a ser estimado é 2.

Para se obter os outros dois valores, foi necessário utilizar os gráficos da ACF (função de autocorrelação) e PACF (Função de autocorrelação parcial), que podem ser visualizados abaixo:

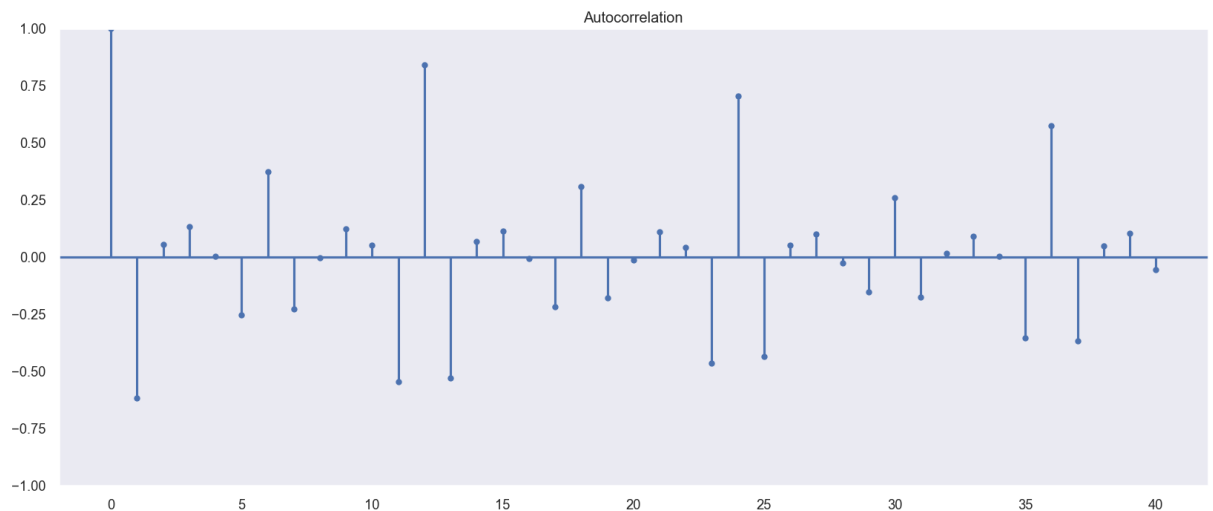


Figura 20 - Gráfico da função de autocorrelação

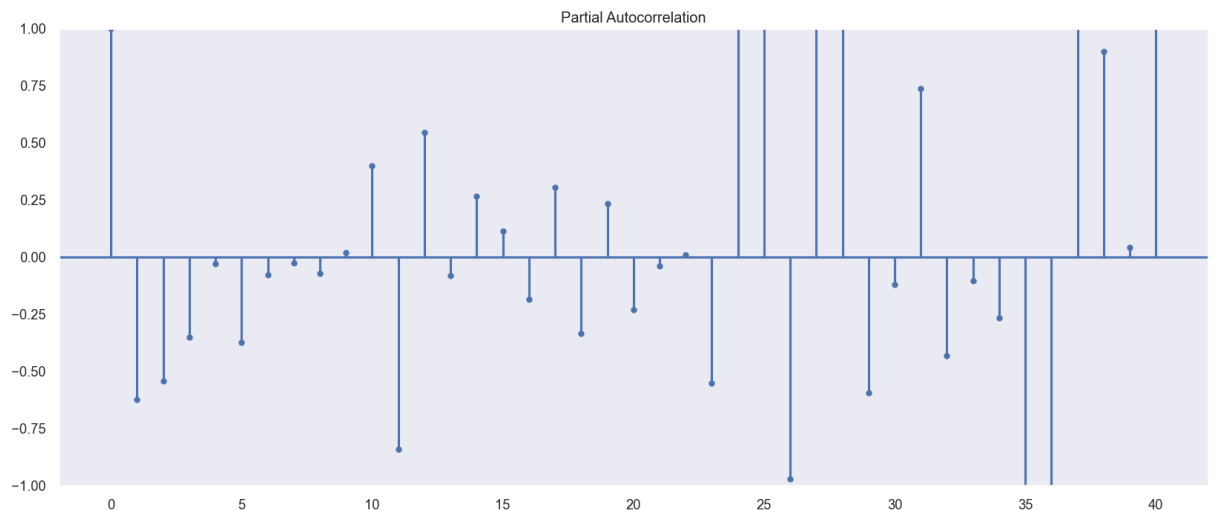


Figura 21 - Gráfico da função de autocorrelação parcial

Para a determinação dos padrões  $p$  e  $q$  foram utilizados os padrões teóricos das ACF e PACF de cada tipo de modelo, foram selecionados para teste os modelos ARIMA (0,2,0), ARIMA(0,2,1), ARIMA(0,2,2) e ARIMA(1,2,1) dado o padrão das funções ACF e PACF encontradas.

### 4.3 Estimação

Nessa etapa, foram estimados os quatro modelos selecionados na etapa anterior (ARIMA 0,1,1) e ARIMA(0,1,0) utilizando a função *ARIMA* da biblioteca *statsmodels*, resultando nos seguintes modelos:

Modelo	AIC (Critério de informação de Akaike)	BIC(Critério de informação Bayesiano)
ARIMA(0,2,0)	2658,40	2660,95
ARIMA(0,2,1)	2683,22	2688,30
ARIMA(0,2,2)	2669,05	2676,68
ARIMA(1,2,1)	2659,09	2666,72

Para prosseguir, foi selecionado o modelo de menor AIC e BIC para que sejam feitos os testes formais em cima dele, resultando em:

SARIMAX Results						
=====						
Dep. Variable:	dessazonalizada		No. Observations:	96		
Model:	ARIMA(0, 2, 0)		Log Likelihood	-1328.202		
Date:	Thu, 30 May 2024		AIC	2658.403		
Time:	16:36:05		BIC	2660.946		
Sample:	01-01-2012		HQIC	2659.430		
	- 12-01-2019					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]
-----						
sigma2	1.075e+11	1.58e+10	6.821	0.000	7.66e+10	1.38e+11
=====						
Ljung-Box (L1) (Q):	30.69	Jarque-Bera (JB):	0.05			
Prob(Q):	0.00	Prob(JB):	0.98			
Heteroskedasticity (H):	0.67	Skew:	0.02			
Prob(H) (two-sided):	0.26	Kurtosis:	2.90			
=====						

Figura 22 - Diagnóstico do modelo ARIMA(0,2,0)



No diagnóstico, foi possível se observar que o modelo passa em todos os testes formais presentes: No teste de Ljung-Box, o modelo apresenta p-valor de 0.00 (valor na parte de baixo em “Prob(Q)”), o que não rejeita a hipótese de que os erros são do tipo ruído branco. No teste de heterocedasticidade, o p-valor de 0.26 (também na parte de baixo em “Prob(H)”), faz com que a hipótese de que os resíduos apresentam variância também não seja rejeitada. Por fim, no teste de Jarque-Bera, foi obtido um p-valor de 0.98, o que não rejeita a hipótese nula de que os dados são normalmente distribuídos. Além dos testes formais, foram feitas análises gráficas também:

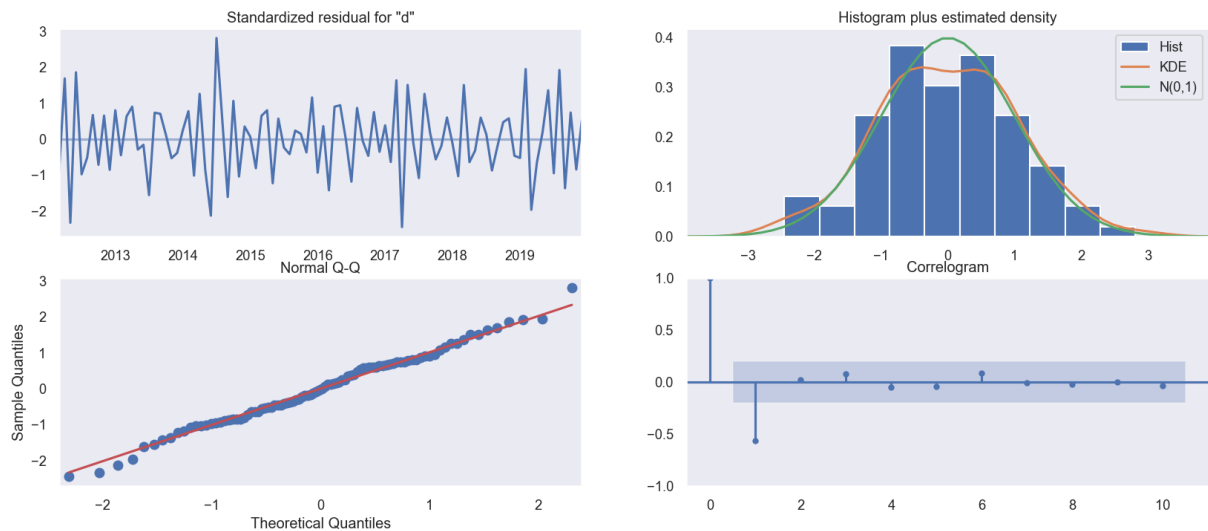


Figura 23 - Testes gráficos do modelo ARIMA(0,2,0)

No gráfico superior esquerdo é possível observar que os resíduos estão aleatoriamente distribuídos e ao redor de zero, sem apresentar uma tendência ou frequência. Já o superior direito mostra que os resíduos estão normalmente distribuídos. O gráfico “Normal Q-Q” também confirma isso, traçando uma linha que representa a distribuição normal (em vermelho) e, por cima, os pontos representando os resíduos, mostrando que os mesmos apresentam praticamente a mesma trajetória. Por fim, o correlograma da direita inferior sugere que não existe autocorrelação entre os resíduos e que eles são efetivamente do tipo ruído branco.

Com isso, é possível prosseguir para o último passo da metodologia Box-Jenkins, que é a previsão utilizando o modelo ARIMA(0,2,0).

#### 4.4 Previsão

Para fazer a previsão, primeiramente foi selecionada a base de dados de 2020 a 2024 e foram aplicada todas as transformações da exata mesma forma que foi feita

com a base de estimação. Nesse sentido, primeiramente a base foi agrupada de forma a se obter os dados mensais do número de passageiros pagantes em voos aéreos no Brasil, ou seja, uma base com 48 meses. A partir disso, foi criada uma previsão para 48 meses com a função *predict* do *statsmodels*. Para se obter os valores reais da previsão foi necessário incluir a sazonalidade nos valores previstos, uma vez que o modelo foi estimado com uma base sem sazonalidade. Para isso, a base de dados real com os 48 meses foi decomposta (assim como feito com a base de estimação) com a decomposição multiplicativa, retornando o valor da sazonalidade para cada mês. Como a predição estava sem sazonalidade, para incluir esse fator, bastou multiplicar a sazonalidade de cada mês pela predição de cada mês, chegando ao resultado real da predição. Com isso, foi criado um gráfico comparando o valor real com o valor predito:

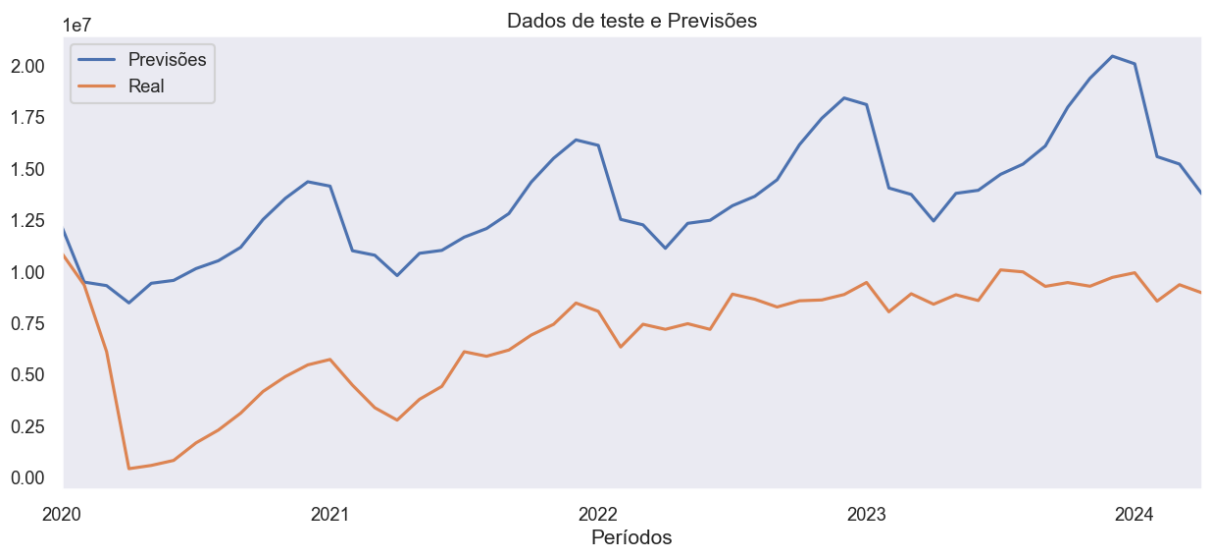


Figura 24 - Valor real vs Valor predito pelo modelo ARIMA

É possível observar o que já era esperado: As previsões mostraram valores muito acima dos valores reais, devido à crise enfrentada pelo setor por causa da pandemia. Além disso, é possível observar que até hoje os valores reais estão significativamente abaixo dos valores da previsão, mostrando que talvez até hoje o mercado não se recuperou. Outro interessante ponto que é possível se observar é que as previsões conseguiram acompanhar os valores reais nas sazonalidades, mostrando picos e vales semelhantes aos valores reais.

## 5 CONCLUSÕES

Nesse projeto de monografia, foram analisados dados da demanda por passagem aérea no Brasil de 2012 a 2024, utilizando tanto da análise gráfica quanto da modelagem de dados para concluir todos os objetivos. Com isso, foi possível mostrar o poder da econometria e da metodologia Box-Jenkins na criação de modelos de séries temporais, extremamente importantes para conseguir entender a relação dos dados ao longo do tempo.

O objetivo geral da monografia foi atendido, visto que, a partir da previsão para os anos de 2020 a 2024 foi possível se observar a diferença do valor predito com o valor real, indicando que talvez o setor de transporte aéreo no Brasil não tenha conseguido se recuperar até hoje, mesmo praticamente 2 anos e meio após as vacinas e o fim do lockdown. Os objetivos específicos também foram atingidos, já que foi feita uma análise em cima da sazonalidade da série de número de passageiros pagantes em voos aéreos, foi possível estimar um modelo estatisticamente significativo e criar previsões com ele e foi possível comparar entender gráfica e numericamente a situação do mercado pré e pós-pandemia.

Portanto, é possível concluir que o estudo concluiu seu papel e respondeu os questionamentos que foram levantados no início de seu desenvolvimento. Como estudos futuros para ampliar o entendimento da problemática, deixo como sugestão a utilização de um SARIMA (Modelos Autoregressivos Integrados de Médias Móveis com Sazonalidade), pois são extremamente úteis para predição de séries que apresentam uma alta sazonalidade (como no caso da demanda por passagens aéreas). Além disso, acredito que um outro estudo para tentar estimar quando o setor se recuperará da crise seria altamente interessante.

### 3 REFERÊNCIAS

OBSERVATÓRIO NACIONAL DE TRANSPORTE E LOGÍSTICA. Boletim de logística – A importância do transporte aéreo para o Brasil. Disponível em: <https://ontl.epi.gov.br/wp-content/uploads/2022/02/Setor-Aereo-Brasileiro-v3.pdf> . Acesso em: 09 de Setembro 2023;

MARCO REGULATÓRIO DA AVIAÇÃO CIVIL - <https://www12.senado.leg.br/publicacoes/estudos-legislativos/tipos-de-estudos/textos-para-discussao/td-42-o-marco-regulatorio-da-aviacao-civil-elementos-para-a-reforma-do-codigo-brasileiro-de-aeronautica> Acesso em: 09 de Setembro de 2023;

BERGIANTE, NISSIA. Análise do Mercado de Transporte Aéreo de Passageiros e dos Modelos de Negócio das Empresas Aéreas;

PAINEL DE INDICADORES DO TRANSPORTE AÉREO DE 2022 <https://www.gov.br/anac/pt-br/assuntos/dados-e-estatisticas/mercado-do-transporte-aereo/painel-de-indicadores-do-transporte-aereo/painel-de-indicadores-do-transporte-aereo-2022> Acesso em: 09 de Setembro de 2023;

PANORAMA 2019 – O setor aéreo em dados e análises – ABEAR: <https://www.abear.com.br/wp-content/uploads/2020/10/Panorama2019.pdf> Acesso em: 09 de Setembro de 2023;

RIBEIRO, LEONARDO. A pandemia da covid-19 e as companhias aéreas brasileiras; INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Ligações aéreas: 2019-2020. Disponível em: <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2101891> . Acesso em: 09 de setembro de 2023;

GUJARATI e PORTER. Econometria básica - 5ª edição 2011;

CORDEIRO. Séries temporais – Análise quantitativa, teoria e aplicações – 1ª edição 2002;

TOLOI, MORETTIN. Análise de séries temporais - 2ª edição 2006.

ANDREA RUSSAR RACHEL (2010) - COMO SE DIFERENCIAM OS INSTITUTOS DA PERMISSÃO, CONCESSÃO, AUTORIZAÇÃO E LICENÇA?. Disponível em: <https://www.jusbrasil.com.br/noticias/como-se-diferenciam-os-institutos-da-permissao-concessao-autorizacao-e-licenca-andrea-russar-rachel/2116750> . Acesso em: 01 de Maio de 2024

Organização da Aviação Civil Internacional (OACI). Disponível em: <https://www.gov.br/anac/pt-br/assuntos/internacional/organismos-internacionais/organizacao-da-aviacao-civil-internacional-oaci>. Acesso em: 01 de Maio de 2024

ANUÁRIO DE TRANSPORTE AÉREO ANAC 2023. Disponível em: <https://www.gov.br/anac/pt-br/assuntos/dados-e-estatisticas/mercado-do-transporte-aereo/panorama-do-mercado/anuario-transporte-aereo>. Acesso em: 01 de Maio de 2024.