

Trabalho de Licenciatura em
Estatística

**Análise da Evolução Temporal de Novos Casos de Infecções
por HIV/SIDA na Província da Zambézia, no (Período de
Janeiro de 2016 à Junho de 2022)**

Autor: Chande Ernesto Jamine

Maputo, Novembro de 2023

FACULDADE DE CIÊNCIAS
Departamento de Matemática e Informática

Trabalho de Licenciatura em
Estatística

**Análise da Evolução Temporal de Novos Casos de Infecções
por HIV/SIDA na Província da Zambézia, no (Período de
Janeiro de 2016 à Junho de 2022)**

Autor: Chande Ernesto Jamine

Supervisor: Zacarias Bernardo Mutombene, MSc., UP

Maputo, Maio de 2024

Dedicatória

Dedico este trabalho à toda família Jamine, e em especial a meus pais, Ernesto Jamine e Herminia Langa pelo apoio incondicional durante todo o processo da formação académica e na elaboração desta monografia, e a minha filha Chaysa Chande Jamine e minha namorada Shelsia Américo Munguambe (Minha Curandeira).

Declaração de Honra

Declaro por minha honra que o presente Trabalho de Licenciatura é resultado da minha investigação e que o processo foi concebido para ser submetido apenas para a obtenção do grau de Licenciado em Estatística, na Faculdade de Ciências da Universidade Eduardo Mondlane.

Maputo, 13 de Novembro de 2023

(Chande Ernesto Jamine)

Agradecimentos

Em primeiro lugar, agradecer a Deus pela graça de vida e protecção em todos os momentos.

À minha família, pelo apoio ao longo dos meus estudos, aos que me acompanharam, aconselharam e incentivaram em todos os momentos dos meus estudos;

Ao Mestre Zacarias Bernardo Mutombene meu supervisor que no momento da preparação da minha tese esteve sempre ao meu lado dando todo tipo de apoio;

À todo corpo docente do DMI, muito em particular aos docentes do Curso de Licenciatura em Estatística, pela preocupação com o sucesso académico;

À Direcção Provincial de Saúde da Província da Zambézia, instituição que me forneceu a base de dados para o estudo;

À Direcção Tecnologias e Informação do Ministério da Justiça no momento da preparação da minha tese;

Aos colegas Jorge Bengala, Arnelio Manhique, Kevin Nhamtumbo, Stélío Mucavele e amigos da turma de 2017 do curso de Licenciatura em Estatística com os quais compartilhei momentos de alegria e amadurecimento.

Resumo

A epidemia do HIV/SIDA, representa um fenómeno global, dinâmico e instável, cuja forma de ocorrência nas diferentes regiões do mundo depende, entre outros determinantes, do comportamento individual e do colectivo. No presente trabalho são apresentados análises que descrevem o comportamento dos novos casos de infecção por HIV/SIDA na Província da Zambézia, situada na zona Centro de Moçambique, classificada como uma das Províncias mais povoadas do País. Este trabalho, tem como objectivo geral analisar a série de casos novos de infecção por HIV na Província da Zambézia, no período de Janeiro de 2016 ao Junho de 2022. Os dados para a realização deste estudo são secundários, correspondentes aos novos casos de infecção por HIV/SIDA mensais em cada uma das portas de entrada das unidades sanitárias da província e comunidade, cujo a base de dados foi fornecida pela Direcção Provincial de Saúde da Zambézia. Para alcançar os objectivos propostos, foi aplicada a metodologia de Box-Jenkins, cujos resultados mostram que a série é estacionária em nível, apesar de apresentar flutuações em alguns meses. No entanto o melhor modelo seleccionado é ARMA (0,1,0) visto que evidenciou melhor desempenho estatístico em relação aos modelos concorrentes e foi usado posteriormente para fazer a previsão de casos futuros de infecção por HIV/SIDA na Província da Zambézia.

Palavras chaves: Infecções, HIV/SIDA, Séries temporais, Metodologia Box & Jenkin

Abstract

The HIV/AIDS epidemic represents a global, dynamic and unstable phenomenon, whose form of occurrence in different regions of the world depends, among other determinants, on individual and collective behavior. This work presents analyzes that describe the behavior of new cases of HIV/AIDS infection in the Province of Zambézia, located in the Central region of Mozambique, classified as one of the most populated Provinces in the country. This work has the general objective of analysing the series of new cases of HIV infection in the Province of Zambézia, from January 2016 to June 2022. The data for carrying out this study are secondary, corresponding to new monthly cases of HIV/AIDS infection in each of the ports entry point of health units in the province and community, whose database was provided by the Provincial Health Directorate of Zambézia. To achieve the proposed objectives, the Box-Jenkins methodology was applied, the results of which show that the series is stationary in level, despite presenting fluctuations in some months. However, the best model selected is ARMA (0,1,0) as it demonstrated better statistical performance in relation to competing models and was subsequently used to predict future cases of HIV/AIDS infection in the Province of Zambézia.

Keywords: Infections, HIV/AIDS, Time series, Box & Jenkin Methodology

Abreviaturas

ADF	Dickey e Fuller Aumentado;
AIC	Critério de Informação Akaike;
AR	Auto regressivo;
ARMA	Modelos Auto Regressivo e de Médias Móveis;
ARIMA	Auto regressive Integral Moving Average;
CNCS	Conselho Nacional de Combate ao SIDA;
DF	Dickey e Fuller;
DPSZ	Direcção Provincial de Saúde da Zambézia;
FAC	Função de Auto Correlação;
FACP	Função de Auto Correlação Parcial;
HIV	Vírus de Imunodeficiência Humana;
KPSS	Kwiatkowski-Phillips-Schmidt-Shin;
MISAU	Ministério da Saúde;
ONUSIDA	Organização das Nações Unidas para o SIDA;
NPCS	Núcleo Provincial de Combate ao SIDA;
PEPFAR	Plano Emergente do Presidente dos Estados Unidos para o Alívio do HIV;
PE	Processo Estocástico;
PP	Phillip Perron;
SARIMA	Modelos sazonais auto-regressivos e de médias móveis;
SIDA	Síndrome de Imunodeficiência Adquirida.

Índice

Introdução	1
1.1. Contextualização.....	1
1.2. Definição do problema.....	2
1.3. Objectivos	3
1.3.1. Objectivo Geral.....	3
1.3.2. Objectivos Específicos	3
1.4. Justificação.....	3
Revisão de Literatura.....	4
2.1. Epidemia do SIDA.....	4
2.2. HIV/SIDA em Moçambique	4
2.3. Série temporais	6
2.3.1. Componentes da série temporal.....	6
2.3.2. Estacionariedade	7
2.3.3. Função de autocorrelação	8
2.3.4. Função de autocorrelação parcial (FACP).....	9
2.4. Teste de Raiz unitária	9
2.5. Modelos de séries temporais.....	11
2.5.1. Modelos auto-regressivos (AR).....	11
2.5.2. Modelos de Média Móvel (MA).....	13
2.5.3. Modelos Mistos Autoregressivos de Médias Móveis (ARMA).....	14
2.5.4. Modelos Auto-regressivos Integrados e de Medias Móveis (ARIMA).....	14
2.3.5. Modelos sazonais auto-regressivos e de médias móveis (SARIMA).....	15
Material e Métodos.....	16
3.1. Material.....	16
3.1.1. Fonte de dados	16
3.1.2. Caracterização da Área de Estudo	17
3.2. Métodos	18
3.2.1. Teste de detecção da componente sazonal	18

3.2.2. Verificação das suposições	19
3.2.3. Teste de estacionaridade	20
3.2.4. Metodologia de Box-Jenkins	21
Resultados e Discussão	26
4.1 Resultados	26
4.1.1. Análise descritiva	26
4.1.2. Verificação das suposições.....	27
4.1.3. Análise da Estacionariedade	29
4.1.4. Identificação do modelo	30
4.1.5. Diagnóstico do modelo estimado	31
4.1.6. Previsão	32
Conclusões e Recomendações	34
5.1. Conclusões	34
5.2. Recomendações	35
Referências Bibliográficas	36
Anexos.....	38

Lista de Figuras

Figura 1 - Mapa dos Distritos da Zambézia.....	17
Figura 2 Comportamento da série de casos de infecção por HIV/SIDA.....	27
Figura 3 - Histograma da série de casos de infecção por HIV/SIDA.....	27
Figura 4 - Representação gráfica dos resíduos do modelo	31
Figura 5 – Representação grafica das previsões num intervalo de confiança de 95%	33
Figura 6 - Comportamento das 1 ^{as} diferenças da série de casos de infecção por HIV/SIDA	38
Figura 7 - Comportamento da sazonalidade das 1 ^{as} diferenças de série de casos de infecção por HIV/SIDA.....	38

Lista de Tabelas

Tabela 1. Propriedades das FAC e FACP nos modelos ARIMA	22
Tabela 2. Propriedades das FAC e FACP nos modelos ARIMA	22
Tabela 3. Estatísticas descritivas da série das infecções.....	26
Tabela 4. Teste de normalidade de Kolmogorov-Smirnov.....	28
Tabela 5. Anova para comparação das médias dos casos positivos de HIV	28
Tabela 6. Teste de Homogeniedade das variâncias	29
Tabela 7. Teste de ADF das infecções registradas em nível.....	29
Tabela 8. Teste de ADF das infecções registradas em nível.....	29
Tabela 9. Modelos gerados pela função <code>auto.arima()</code>	30
Tabela 10. Teste de ADF dos resíduos do modelo estimado.....	31
Tabela 11. Teste de Box-Pierce dos resíduos do modelo	31
Tabela 12. Teste de Ljung-Box dos resíduos do modelo.....	32
Tabela 13. Medidas para avaliar a capacidade de prever do modelo estimado.....	32
Tabela 14. Previsão dos casos de infecção por HIV/SIDA	33

Capítulo I

Introdução

Neste capítulo, foram abordados os aspectos principais que levaram ao desenvolvimento desse estudo, desde a contextualização, a definição do problema, os objectivos e a justificação que levou ao desenvolvimento do estudo.

1.1. Contextualização

A epidemia do HIV/SIDA, representa um fenómeno global, dinâmico e instável, cuja forma de ocorrência nas diferentes regiões do mundo depende, entre outros determinantes, do comportamento individual e do colectivo. A Síndrome de imunodeficiência Adquirida (SIDA) destaca-se entre as enfermidades infecciosas emergentes pela grande magnitude e extensão dos danos causados às populações. Desde a sua origem, cada uma das suas características e repercussões tem sido exaustivamente discutida pela comunidade científica e pela sociedade em geral (Brito *et al.*, 2000).

Segundo Brasil (2010), desde o aparecimento dos primeiros casos da Síndrome da imunodeficiência Adquirida no final da década de 1970, a doença tem sido alvo de atenção no plano mundial e seu crescimento exponencial na década de 1980 a tornou problema de Saúde Pública enfrentado até os dias de hoje. O vírus da imunodeficiência humana (HIV) é o agente causador da SIDA sendo que é um vírus linfotrópico com afinidade preferencial para os linfócitos T CD4⁺ (responsáveis, em parte, pelo controlo do sistema imunológico). A SIDA é a fase mais avançada da infecção pelo vírus da imunodeficiência humana.

A epidemia do HIV/SIDA continua a preocupar todo o mundo, estimativas indicam que aproximadamente 37 milhões de pessoas vivem com HIV/SIDA em todo mundo, dos quais 19 milhões (52%) são da África Subsariana e em Moçambique aproximadamente 1,5 milhões (4%) de pessoas vivem com HIV/SIDA em todas as idades, sendo que Moçambique está na lista dos 8 países com as taxas de prevalência mais elevadas em todo mundo (UNAIDS, 2015).

Nos últimos anos Moçambique tem recebido ajuda do PEPFAR¹, USAID², Global Found para apoiar projectos com vista ao alívio/redução de novos casos de infecção de HIV/SIDA, sendo que a Província da Zambézia encontra-se no centro das atenções na zona Centro do País, por ser a segunda mais populosa do País, com a maior parte da população na zona rural com elevados índices de pobreza, com difícil acesso aos cuidados de saúde, educação, informação sobre a doença e com uma taxa de prevalência de 12.6%. Neste contexto, no presente estudo pretende-se analisar a série de casos novos de infecção por HIV na Província da Zambézia no período de Janeiro de 2016 à Junho de 2022, com a finalidade de estimar um modelo para prever os novos casos de infecção por HIV/SIDA na Província da Zambézia para os próximos 12 meses.

1.2. Definição do problema

Moçambique está a registar um crescimento preocupante de casos de HIV/SIDA com um impacto negativo no desenvolvimento socioeconómico do país. Face a esta situação, o MISAU declarou o HIV/SIDA como uma emergência nacional e precisa do envolvimento de todos os sectores da sociedade, a todos níveis, com acções preventivas e de mitigação dos seus efeitos, numa perspectiva integradora e intersectorial de forma a reverter o cenário actual (PNC ITS/HIV/SIDA, 2010). Porque o HIV/SIDA é um problema que preocupa o mundo, os sistemas de saúde pública e privada tiveram suas estruturas ampliadas e novas estratégias elaboradas para tentar reduzir a epidemia, mas ainda assim, novos desafios se impõem diariamente por esta epidemia, porque continuam sendo registadas novas infecções em todo mundo.

De acordo com a ONUSIDA, em todo mundo são notificados cerca de 2.5 milhões de novos casos por ano, sendo que 70% são da África Subsaariana. Segundo Governo de Moçambique (2015), o país foi classificado em terceiro lugar no mundo depois da Nigéria e da África do Sul, no que diz respeito à novas infecções entre 22 países que contribuem para o elevação mundial de novas infecções de HIV. Este país vive uma situação de pandemia generalizada que afecta todos os segmentos da população, ainda que determinados grupos populacionais se destaquem, quer em termos de vulnerabilidade e propensão ao vírus, quer em termos de prevalência e infecção. Segundo PEPFAR, estima-se que aproximadamente 20% das pessoas vivendo com HIV/SIDA em Moçambique estão na Província da Zambézia.

¹ Plano de Emergência do Presidente dos Estados Unidos para o Alívio do SIDA

² Agência dos Estados Unidos para o Desenvolvimento Internacional

Assim sendo, diante dos pontos apresentados envolvendo a pademia da SIDA em Moçambique, aliada a projectos implementados com vista a redução dos índices de prevalência, bem como novas infecções pelo HIV/SIDA e procurar condições para a elevação das taxas de retorno ao consumo anti-retroviral dos que abandonaram o tratamento. Desta forma, para o desenvolvimento deste estudo pergunta-se: Qual é a tendência dos novos casos de infecções pelo HIV/SIDA na província da Zambézia?

1.3. Objectivos

1.3.1. Objectivo Geral

Analisar a série de casos novos de infecção por HIV na Província da Zambézia no período de Janeiro de 2016 à Junho de 2022.

1.3.2. Objectivos Específicos

- Descrever o comportamento da série de casos novos de infecção por HIV/SIDA na Província da Zambézia;
- Estimar o melhor modelo que gera a série de novos casos de infecção por HIV/SIDA na Província da Zambézia;
- Prever os novos casos de infecção por HIV/SIDA na Província da Zambézia para os próximos 12 meses;
- Monitorar a previsão de novos casos de infecção por HIV/SIDA na Província da Zambézia.

1.4. Justificação

O estudo da evolução temporal da epidemia do HIV/SIDA na Província da Zambézia, evidencia-se como um campo de estudo muito importante, porque permitirá a visualização do comportamento da incidência dessa doença, além de mostrar o impacto das acções dos serviços de saúde locais, dos programas implementados com vista a redução da epidemia com intervenções baseadas em evidências rumo à uma geração livre da SIDA, incluindo Testagem e Aconselhamento ao HIV, Circuncisão Médica Masculina Voluntária, Tratamento Anti-retroviral, cuidados clínicos e de base comunitária para as pessoas vivendo com HIV/SIDA através do fundo disponibilizado pelos doadores e consequentemente apontar a necessidade de desenhar políticas públicas específicas que sejam efectivas na conscientização da população em relação ao comportamento de exposição ao risco de contrair HIV/SIDA bem como advogar informações sobre os seus métodos de prevenção.

Capítulo II

Revisão de Literatura

Neste capítulo foi apresentado de forma clara a situação actual do HIV/SIDA em Moçambique. Faz-se ainda, uma apresentação dos principais conceitos e a descrição das técnicas estatísticas usadas para o desenvolvimento do estudo.

2.1 Epidemia do SIDA

A SIDA é uma doença grave, quase sempre fatal e relativamente recente para a qual não existe cura. A descrição original da SIDA ocorreu em 1981 nos Estados Unidos da América (EUA) em jovens homossexuais do sexo masculino, 26 dos quais com sarcoma de Kaposi e 5 com Candidíase oral e Pneumonia por *pneumocystis* (Almeida *et al.*, 2005).

Vinte anos depois de ser descoberta, a SIDA tornou-se a doença mais devastadora da história, tendo infectado mais de 60 milhões de pessoas em todo mundo em que cerca de um terço (20 milhões) já faleceram (Portugal INCS, 2002).

Já em 1985 um Comité de Especialistas da Organização Mundial da Saúde (OMS), reunido em Banguí, tinha feito a definição do primeiro caso clínico de SIDA em África. A OMS passou então a recomendar aos Países membros, a constituição de comités nacionais de luta contra o HIV/SIDA.

A África Subsaariana continua a ser a região do mundo mais afectada pela epidemia. Existem cerca de 37 milhões de pessoas vivendo com HIV em todo mundo, dos quais 52% de todas pessoas vivendo com HIV encontram-se nesta região. De todas regiões africanas, a África Austral é a mais afectada com elevados índices de prevalência (UNAIDS, 2015).

2.2 HIV/SIDA em Moçambique

Segundo o Conselho Nacional de Combate ao HIV/SIDA (CNCS, 2004), o primeiro caso de SIDA em Moçambique foi diagnosticado em 1986. Tratava-se de um cidadão estrangeiro que já vinha infectado quando entrou no País. Em Agosto de 1986 foi criado o primeiro organismo de combate ao SIDA com a designação de Comissão Nacional do SIDA, actualmente designado por Comissão Nacional de Combate ao SIDA.

Em 1987, durante vários meses foi levado a cabo em várias cidades do País, um inquérito sero-epidemiológico, que detectou infecções do HIV. Entretanto no decurso do mesmo ano são notificados os primeiros 5 casos clínicos em cidadãos nacionais. No período que segue, até 1989, o número de casos duplica em cada ano, tendo em Julho de 1989 chegado ao total de 41. O grupo etário mais atingido era dos 20 a 29 anos (CNCS, 2004).

A situação de guerra em que o país vivia, levou ao deslocamento de povoações em busca de refúgio e segurança, quer noutras áreas no País, quer em Países vizinhos onde as taxas de prevalência do HIV já eram elevadas. O governo foi obrigado a investir maior parte dos recursos disponíveis no esforço da defesa. Por força dessas dificuldades conjunturais, o combate ao HIV/SIDA passou a ser aparentemente um problema menos importante para o País.

Portanto, durante a década 90 o conhecimento sobre a evolução do HIV/SIDA não progrediu em Moçambique. As pessoas continuaram a infectar-se inocentemente pois não havia informação sobre os riscos de sexo sem protecção e outras de propagação do HIV/SIDA.

Em 1993 e 1996 foram feitos estudos de Conhecimento Atitude e Prática (CAP), a estudantes da UEM, que foram apresentados nas Jornadas de Saúde de 1994 e 1997. Esses estudos mostram resultados alarmantes e com agravante de que tinha havido pouca evolução entre 1993 e 1997 no que concerne a divulgação da informação e combate a esta pandemia.

Entretanto, pensava-se que o aumento das taxas de prevalência era devido ao regresso maciço de populações que em resultado da guerra tinham-se refugiado nos Países vizinhos, onde as taxas de sero-prevalência eram altas. Desta forma, acreditava-se que com o reassentamento tudo iria se normalizar.

Em 1999, o Governo decidiu actuar com maior energia no combate a epidemia e criou um Grupo Técnico de Apoio a luta contra o HIV/SIDA, que compreendia, além do MISAU, outras estruturas de investigação, planificação, divulgação e combate, envolvendo todas as estruturas da sociedade moçambicana, e todos possíveis parceiros internacionais.

Em Maio de 2001, criou-se em Maputo os primeiros Gabinetes de Aconselhamento e Testagem Voluntária do HIV/SIDA (GATV), com vista a acelerar o acesso do diagnóstico da infecção pelo HIV com oferta de uma testagem rápida e desburocratizada. Estes gabinetes surgem como uma estratégia do MISAU mais vinculada à prevenção, enquadrada na resposta nacional do HIV/SIDA.

2.3. Série temporais

Segundo Sousa (2006), uma série temporal é qualquer conjunto X de observações ordenadas no tempo, $X = X_t, t = 1, 2, \dots, N$ onde t é o índice de tempo e N é o número de observações.

Uma série temporal é uma observação de um processo estocástico. Suponha então, uma amostra de valores observados de tamanho T da variável aleatória X_t , tal que: $\{X_1, X_2, \dots, X_T\}$ e considere uma colecção de valores independentes e identicamente distribuídos ε_t tal que: $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T\}$ e ε_t tem distribuição normal com média 0 (zero) e variância σ^2 ($\varepsilon_t \sim N(0, \sigma^2)$), então ε_t é considerado um processo Gaussiano com ruído branco. A sequência de valores observados, X_t , representa uma série temporal. Desta forma, uma série temporal é uma realização de um processo estocástico (Morettin & Toloi, 2004).

2.3.1. Componentes da série temporal

Segundo Morettin e Toloi (2004), uma série temporal geralmente é, composta por quatro componentes:

- **Tendência:** A tendência indica a direcção geral dos valores estudados. A principal característica de componente é o movimento suave registado num período longo de tempo, direccionando os dados de modo constante, crescente ou decrescente.
- **Ciclo:** São as variações irregulares ou alterações na demanda passada resultante de factores excepcionais, como greves ou catástrofes climáticas, que não podem ser previstos e, portanto, incluídos no modelo.
- **Sazonalidade:** As variações sazonais são flutuações que se repetem periodicamente, acompanhando um padrão temporal relacionado ao factor tempo. A sazonalidade refere-se às mudanças ou variações cíclicas de curto prazo. São bastante parecidas com os cíclicos, com diferença de que os fenómenos cíclicos são caracterizados por variações que oscilam em torno da tendência a intervalos aproximadamente regulares de tempo e em longo prazo.
- **Erro (Aleatoriedade):** é tudo aquilo que não é explicado pelas outras componentes da série ou seja, é o que o modelo estimado não consegue captar. Um erro é dito ruído branco quando possui distribuição normal com a média zero e variância constante. Quando os erros não são correlacionados significa que o modelo elaborado conseguiu explicar o máximo da série de dados.

O estudo de uma série temporal, consiste na descrição das componentes que cada série temporal possui. Normalmente desconsidera-se a componente cíclica, devido as suas características e considera-se que a componente aleatória existente em todas as séries temporais, devido ao facto de ser praticamente impossível a sua eliminação. A principal utilização de uma análise de uma série temporal é a previsão de movimentos futuros, como tentativa de controlo e de planificação, (Johnston e Dinardo, 2001).

2.3.2. Estacionariedade

Um dos pressupostos fundamentais para a previsão de uma série temporal a longo prazo é o pressuposto de estacionaridade. Segundo Gujarati (2006), um processo estocástico é estacionário quando a sua média e a sua variância são constantes ao longo do tempo e quando o valor da covariância entre dois períodos de tempo depende apenas da distância, do intervalo ou da defasagem entre os dois períodos de tempo, e não do próprio tempo em que a covariância é calculada.

Na literatura sobre séries temporais, tal processo estocástico é conhecido como fracamente estacionário, ou estacionário em covariâncias, ou ainda estacionário de segunda ordem. E uma série temporal não-estacionária terá uma média que varia com o tempo ou uma variância que varia com o tempo, ou as duas situações.

Segundo Gujarati (2006) uma série temporal estocástica possui as seguintes propriedades:

$$\textbf{Média:} \quad E(Y_t) = \mu \quad (1)$$

$$\textbf{Variância:} \quad \text{var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2 \quad (2)$$

$$\textbf{Covariância:} \quad \gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)] \quad (3)$$

Onde: γ_k é a covariância (ou autocovariância) na defasagem k , é a covariância entre os valores de γ_t e γ_{t+k} , ou seja entre dois valores de Y separados por k períodos. Se $k = 0$, obtemos γ_0 , que é simplesmente a variância de γ ($=\sigma^2$); se $k = 1$, γ_1 é a covariância entre dois valores adjacentes de γ , o tipo de covariância que discute a autocorrelação.

Se uma série temporal não for estacionária é chamada de série temporal não-estacionária (quando se trata de estacionaridade fraca) porque as vezes, a não estacionaridade pode ser devido a uma mudança na média (Gujarati, 2006).

Geralmente quando se trabalha com séries temporais não-estacionárias em modelos econométricos, estes costumam gerar regressões espúrias³. Isso é visível através de coeficiente de determinação (R^2). Esse problema ocorre quando as séries temporais envolvidas exibirem fortes tendências (movimentos ascendentes ou descendentes continuados), o R^2 apresenta um valor muito alto, e isso se deve a presença da tendência e não a verdadeira relação entre as variáveis. Pelo facto dos resíduos de uma regressão espúria serem não-estacionários, os teste F e t, não são validos. Portanto, torna-se necessário a realização de testes para verificar se um processo estocástico é estacionário, dado que as estimativas dos modelos que utilizam dados no tempo dependem da estacionaridade das séries.

Uma estatística importante na análise de séries temporais é o coeficiente de autocorrelação ρ . A autocorrelação é usada para descrever a correlação entre dois valores da mesma série temporal, em diferentes períodos de tempo existente.

2.3.3. Função de autocorrelação

Define-se a função de autocovariância de uma série temporal estacionária (X_t) como:

$$\gamma_x(h) = \text{cor}(X_{t+h}, X_t) \quad (4)$$

A função de autocorrelação (FAC) de uma série temporal estacionária (X_t) é definida como

$$\rho_x(h) = \frac{\gamma_x(h)}{\gamma_x(0)} = \text{cor}(X_{t+h}, X_t) \quad (5)$$

As funções de autocovariância e de autocorrelação fornecem uma medida útil do grau de dependência entre os valores de uma série temporal em diferentes períodos. As autocorrelações medem ainda o tamanho e a força da “memória” do processo.

O gráfico das autocorrelações amostrais *versus* h e chamado de correlograma. Tal gráfico apresenta valores que serão utilizados para caracterizar as populações lineares ou não do mecanismo gerador do processo (Granger & Newbold, 1986). Porém, não é simples examinar um correlograma e extrair dele as correspondentes propriedades populacionais. O que é necessário se fazer é averiguar alguns modelos plausíveis que provejam correlogramas de formas relacionadas.

³ Uma **regressão espúria** é uma relação estatística entre duas variáveis, mas onde não existe nenhuma relação de causa-efeito entre elas.

As funções de autocovariância e autocorrelação amostrais podem ser calculadas para qualquer conjunto de dados e não estão restritas a observações de séries temporais estacionárias (Brockwell & Davis, 1996). Para dados contendo tendência, a FAC exibirá um decaimento lento a medida que t aumenta. Assim, o correlograma pode ser utilizado como indicador da estacionariedade ou não de uma série temporal (Brockwell & Davis, 1996).

2.3.4. Função de autocorrelação parcial (FACP)

A função de autocorrelação parcial (FACP) é definida como a sequência de correlações entre $(X_t \text{ e } X_{t-1})$, $(X_t \text{ e } X_{t-2})$, $(X_t \text{ e } X_{t-3})$ e assim por diante, desde que os efeitos de defasagens anteriores sobre X_t permaneçam constantes (Hill, Griffiths e Judge, 2010). A FACP é calculada como o valor do coeficiente Φ_{kk} na equação

$$X_t = \Phi_{k1}X_{t-1} + \Phi_{k2}X_{t-2} + \dots + \Phi_{kk}X_{t-k} + e_i \quad (6)$$

A k -ésima autocorrelação parcial é dada pelo coeficiente Φ_k . Uma característica importante da função de autocorrelação parcial é que leva em consideração toda a história do processo até ao desfasamento de ordem k correspondente, ou seja, mede a correlação adicional entre X_t e X_{t+k} , uma vez controlados os efeitos de todos os desfasamentos intermédios. Um caso particular de um processo fracamente estacionário é o processo ruído branco⁴. Num processo deste tipo a média, a co-variância são nulos e a variância é finita: $\mu = \gamma_k = 0$ ($k \neq 0$) e $\gamma_0 = \sigma^2$ (Enders, 2004).

2.4. Teste de Raiz unitária

Segundo Franco (2004), a eficiência das estimações dos modelos que utilizam séries temporais depende da estacionariedade da série, são descritos alguns testes capazes de detectar se o processo estocástico é ou não estacionário, conhecidos como testes de raiz unitária.

Esses testes são capazes de detectar se a série foi suficientemente diferenciada para se tornar estacionária. Ainda segundo Franco (2004), se uma série deve ser diferenciada d vezes de tornar-se estacionária, então ela contém d raízes unitárias e é dita ser integrada de ordem d , denotada por $I(d)$.

⁴ O processo de ruído branco ocorre quando a série temporal não depende do seu passado, nesta situação a série não é autoprojectável e todas as autocorrelações são nulas.

Segundo Gujarati (2006), o correlograma de um processo puramente de ruído branco⁵ apresenta coeficientes de autocorrelação em torno de zero em várias defasagens, ilustrando a imagem de um correlograma de uma série temporal estacionária. Deste modo, se o correlograma de uma série temporal efectiva se parece ao correlograma de uma série temporal de ruído branco, pode-se dizer que essa série temporal provavelmente é estacionária. Portanto, essa percepção intuitiva que é a análise gráfica é o ponto de partida dos testes de estacionaridade mais formais. Para testar a série temporal se é integrada ou não é necessário realizar teste de raiz unitária e o presente estudo fez menção de três testes:

a) Dickey-Fuller e Dickey-Fuller Aumentado

Segundo Gujarati(2006), o teste de Dickey-Fuller (DF) e Dickey-Fuller Aumentado (ADF) é um meio de diagnóstico para testar se a série temporal é integrada ou não. Onde a hipótese nula do teste de DF é a existência de raiz unitária, onde o termo ε_i do erro estatístico segue hipóteses clássicas, tem média zero, variância constante e a série possui ruído branco. Neste teste, se o valor absoluto calculado da estatística t for maior que os críticos absolutos t de DF ou Mackinnon-DF não se rejeitam a hipótese de que a série temporal seja estacionária caso contrario, série é não estacionária.

O teste de Dickey-Fuller tem alguma desvantagem, pois, na realização deste teste assume-se que o termo de erro ε_i não apresenta autocorrelação serial, não incorporado o caso em que os resíduos apresentam correlação serial. Para ultrapassar esse problema Dickey e Fuller desenvolveram um teste conhecido por teste de Dickey-Fuller Ampliado ou aumentado (ADF), que consiste em adicionar nas equações descritas no teste DF os valores desfasados na variável dependente ΔX_t , após a inclusão dos termos desfasados da variável dependente.

b) Phillips e Perron (PP)

O teste de Phillips e Perron (PP) é baseada na abordagem semiparamétrica a partir do princípio da invariância e dos movimentos Brownianos, considerando que os termos de perturbação estatísticos são realmente correlacionados. Assim como a maioria dos testes clássico, para estacionariedade o procedimento PP possui como hipótese nula a existência de uma raiz unitária (Alencar, 2010).

⁵ Uma série é puramente aleatório (*ruído branco*) quando seus resíduos tem média zero, a variância σ^2 é constante e é não serial correlacionado com o tempo

O teste de PP apresenta algumas vantagens em relação aos testes de DF e ADF, dado que, este é robusto, as formas de heteroscedasticidade dos resíduos e durante a realização não é necessário especificar o número de defasamentos na estimação das regressões descritas acima.

c) Kwiatkowski-Phillips-Schmidt-Shin (KPSS)

O teste de Kwiatkowski-Phillips-Schmidt-Shin (KPSS), também é baseado na abordagem semiparamétrica a partir do princípio da invariância e dos funcionais movimentos Browianos. A sua hipótese nula é a inexistência de raiz unitária (Alencar e Marques, 2007).

Se o valor absoluto da estatística KPSS for menor que os críticos absolutos de KPSS a 5% não se rejeitam a hipótese de que a série temporal seja estacionária caso contrário, a série é não estacionária.

2.5. Modelos de séries temporais

Para aplicação das séries temporais é necessário primeiro fazer uma investigação sobre as características das mesmas. Esta investigação consiste em analisar as estatísticas descritivas dos dados que compõem a série, a função de autocorrelação, bem como verificar as questões de estacionaridade e a necessidade ou não de transformações. Isto se faz pela necessidade de aplicação da maioria dos métodos (Gujarati, 2006).

2.5.1. Modelos auto-regressivos (AR)

Segundo Gujarati (2006), se uma série temporal qualquer for estacionária, esta pode ser modelada de vários modos. A característica principal do modelo auto-regressivo é que no lugar das variáveis independentes, o processo vai utilizar os valores prévios da série temporal para a estimação do modelo. No modelo auto-regressivo a série das observações históricas Z_t é descrita por seu valor

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \dots + \Phi_p X_{t-p} + \varepsilon_t \quad (7)$$

passado regredido e pelo ruído aleatório. Assim o modelo de ordem p , AR (p) é dado por:

Onde: Φ_i é o parâmetro a ser estimado, que descreve como o valor corrente de X_t se relaciona com os valores passados de X_{t-i} , para $i = 1, 2, 3, \dots, p$. Para simplificar a representação do modelo, utiliza-se um operador auto-regressivo de ordem p , dado por:

$$\Phi(B) = 1 - \Phi_1 B - 1 - \Phi_2 B^2 - \dots - \Phi_p B^p \quad (8)$$

E reduz-se o modelo auto-regressivo de ordem p para:

$$\Phi(B) = X_t = \varepsilon_t \quad (9)$$

Segundo Werner & Ribeiro (2003), o modelo auto-regressivo de ordem 1 ou AR(1) é a versão mais simples dessa classe de modelos. Sua representação algébrica é dada por:

$$X_t = \Phi_1 X_{t-1} + \varepsilon_t \quad (10)$$

Para processos auto-regressivos de primeira ordem, AR(1), também chamados de processos Markovianos, a estacionariedade é obtida se $|\Phi| < 1$ e as autocovariâncias (γ_k) sejam independentes (Chatfield, 2004). No caso do modelo AR(1), as autocovariâncias são dadas por:

$$\gamma_k = \Phi_1^k \gamma_0 \quad (11)$$

As autocorrelações ρ_k são dadas pela equação:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \Phi_1^k, \quad \text{para } k = 0, 1, 2, \dots, p \quad (12)$$

Segundo Liebel (2004), os processoa de segunda ordem, AR(2), têm sua estacionariedade assegurada quando:

$$\Phi_1 + \Phi_2 < 1, \Phi_2 - \Phi_1 < 1 \text{ e } |\Phi| < 1$$

A função de autocorrelação decai exponencialmente quando Φ_1 é positivo, quando Φ_1 é negativo, a função de autocorrelação também decai exponencialmente, mas apresenta alternância de sinais positivos ou negativos Werner & Ribeiro (2003).

2.5.2. Modelos de Média Móvel (MA)

De acordo com Johnstor & Dinardo (2001), o modelo de médias móveis de ordem q , ou seja MA (q), descreve um conjunto de dados como sendo resultado da combinação linear dos choques aleatórios que se comportam como ruído branco, ocorridos no período corrente e nos períodos passados, isto é:

Alternativamente, () pode ser escrito como:

$$Y_t = v + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q} = v + \sum_{i=1}^q \alpha_i \varepsilon_{t-i} \quad (13)$$

$$Y_t = v + \beta(B) \varepsilon_t, \text{ com} \quad (14)$$

$$\beta(B) = 1 + \beta_1 B + \beta_2 B^2 + \dots + \beta_q B^q \quad (15)$$

Onde: B^q a representar o polinómio do operador atraso.

Segundo Barros (2005), a condição de estacionariedade em processos MA (q) é sempre satisfeita e portanto em geral é desejável impor restrições para que ele satisfaça uma condição chamada invertibilidade semelhante à condição de estacionariedade.

Um processo MA (q) é invertível se a equação $\theta(B) = 0$ tem todas as raízes fora do círculo unitário, isto é, se $\theta(B) \neq 0$ para $|B| \leq 1$. O processo MA (1), é expresso pela seguinte equação:

$$Y_t = v + \varepsilon_t + \beta_1 \varepsilon_{t-1} \quad (16)$$

É caso particular de MA (q) quando $q = 1$. Caso $|\beta_1| < 1$ o processo é dito estacionário e invertível. As autocorrelações são obtidas como:

$$\rho_k = -\frac{\beta}{(1 + \beta^2)} \quad (17)$$

A FAC é zero para defasagens maiores que q e a FACP apresenta uma combinação de decaimentos exponenciais e sinusoidais.

2.5.3. Modelos Mistos Autoregressivos de Médias Móveis (ARMA)

Os modelos auto-regressivos e de médias móveis é a junção dos modelos AR e MA. Denotados por ARMA(p,q) um processo auto-regressivo e de médias móveis de ordem (p,q). Em alguns casos, pode ser necessário utilizar um grande número de parâmetros em modelos puramente AR ou puramente MA. Nesses casos, é vantajoso misturar os componentes de um modelo AR com os componentes de um modelo MA, gerando assim, um modelo ARMA. O modelo ARMA (p,q) exigirá um número menor de termos e pode ser expresso conforme a seguinte equação:

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \dots + \Phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (18)$$

O modelo ARMA mais simples ou frequentemente utilizado é o ARMA (1; 1)

$$\rho_1 = \frac{(1 - \Phi_1 \theta_1)(\Phi_1 \theta_1)}{1 + \theta_1^2 + \Phi_1 \theta_1} \quad (19)$$

Onde: $\rho_k = \Phi_1 \rho_{k-1}$, para $K > 1$.

A função de autocorrelação do modelo ARMA (p; q) apresenta características da função MA (q) para as defasagens $k < q$, pelo facto de a memória do componente de médias móveis durar apenas q períodos. Nas defasagens maiores que $k+1$ as características são iguais às de um modelo AR (p).

2.5.4. Modelos Auto-regressivos Integrados e de Medias Móveis (ARIMA)

Segundo Luz et al. (2005), os modelos AR, MA e ARMA, são apropriados para descrever séries temporais estacionárias, ou seja, que se desenvolvem no tempo em torno de uma média constante (séries estacionárias). Assim, para ajustar modelos a séries temporais não-estacionárias é necessário remover as fontes de variação não-estacionárias, desde que as mesmas não tenham um comportamento explosivo. A remoção da tendência consiste em tomar um número finito de diferenças d, da série temporal original. A série que se torna estacionária através deste procedimento é chamada de não-estacionária homogênea.

Denotando a série diferenciada por:

$$X_t = \Delta^d Y_t = (1 - B)^d Y_t \quad (20)$$

O processo autoregressivo integrado e de médias móveis denotado ARIMA (p, d, q) onde **p** é a ordem da componente autoregressiva, **d** é o numero de diferenças tomadas na série e **q** é a ordem da componente de médias móveis é dado por:

$$X_t = \alpha + \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q} \quad (21)$$

ou de forma equivalente,

$$\alpha(B) (1 - B)^d Y_t = \beta(B) \varepsilon_t \quad (22)$$

Onde: ε_t é ruído branco com média zero, $\Phi(B)$ e $\beta(B)$ são respectivamente os polinómios autoregressivos e de medias móveis.

Portanto, pode-se descrever todos os modelos vistos anteriormente utilizando a nomenclatura ARIMA, isto é:

- i) ARIMA(p,0,0) = AR(p);
- ii) ARIMA(0,0,q) = MA(q);
- iii) ARIMA(p,0,q) = ARMA(p,q).

2.3.5. Modelos sazonais auto-regressivos e de médias móveis (SARIMA)

Segundo Pellegrini & Fogliato (2001), a maior parte das séries temporais, exhibe uma característica periódica que se repete a cada s intervalos de tempo. Por exemplo, em séries compostas por observações mensais e sazonalidade anual, s é igual a 12. Define-se,

$$\Delta_s = (1 - B^s) \quad (23)$$

como operador da diferença sazonal. Por conseguinte:

$$\Delta_s X_t = (1 - B^s) X_t = X_t - X_{t-s} \quad (24)$$

é a primeira diferenciação sazonal. De um modelo geral, D diferenciações sazonais podem ser requeridas para produzir uma série estacionária. Para Liebel (2004), nos casos em que o processo apresenta componentes sazonais, deve-se acrescentar essa informação no modelo, que terá a seguinte representação:

$$\Phi(B^S) \nabla_S^D X_t = \theta(B) \alpha_t \quad (25)$$

De forma resumida obtém-se a formula genérica do modelo autoregressivo integrado e de media móvel sazonal multiplicativo de ordem $(p, d, q) \times (P, D, Q)_S$ – SARIMA $(p, d, q) \times (P, D, Q)_S$, dada por:

$$\Phi_p(B) \Phi_p(B^S) \nabla^d \nabla_S^D X_t = \theta_q(B) \theta_q(B^S) \varepsilon_t \quad (26)$$

μ_i

Capítulo III

Material e Métodos

Neste apresenta-se a fonte de dados, os softwares usados para a aplicação dos modelos estatísticos e são apresentados os principais passos para a obtenção dos resultados da investigação. O estudo foi conduzido com o auxílio dos *softwares* estatísticos Eviews versão 11 e o pacote de programação RStudio na versão 4.2.1. Para sobre os testes estatísticos feitos, considerou-se o nível de significância de 5%.

3.1 Material

3.1.1. Fonte de dados

Para a realização do trabalho, foi usada uma base de dados secundária, disponibilizada pela Direcção Provincial de Saúde da Província da Zambézia (DPSPZ), contendo observações referentes aos novos casos de infecção de HIV/SIDA notificados no período de Janeiro de 2010 à Junho de 2016 em todos os distritos da Província.

Estes casos, são registados diariamente ao nível das unidades sanitárias, na comunidade e resumidas mensalmente pelos distritos, depois compilados a nível da Direcção Provincial da Saúde da Zambézia (DPSZ) e posteriormente encaminhados ao Ministério da Saúde (MISAU) através das diferentes plataformas (SISMA, Módulo Básico e DHIS2).

3.1.2. Caracterização da Área de Estudo

Zambézia é uma província no centro de Moçambique, tem uma área de 103 127 km² e uma população de 3 792 000 habitantes de acordo com os resultados do Censo de 2007, representando um aumento de 31,1% dos habitantes em apenas dez anos comparativamente ao Censo de 1997. A capital da Província é a Cidade de Quelimane, que dista cerca de 1600 km ao Norte de Maputo, que é a capital do país.

A Província da Zambézia está limitada ao norte pelas províncias de Nampula e Niassa, ao leste pelo Canal de Moçambique, ao sul pela província de Sofala e ao oeste pela província de Tete e pelo Malawi.

O topónimo Zambézia foi criado em 1858 por decreto régio, abrangendo as capitanias de Quelimane e “Rios de Sena” (o Zambeze). O que hoje chamamos Província da Zambézia foi durante muito tempo o “distrito de Quelimane”, criado em 1817, extinto e incorporado no de Sena em 1829 e reposto em 1853.



Figura 1 - Mapa dos Distritos da Zambézia

3.2. Métodos

Foram por coseguite, apresentados todos os passos seguidos para a concretização dos objectivos delimitados no início do estudo.

3.2.1. Teste de detecção da componente sazonal

Para a investigação da componente sazonal realizou-se o teste F de Snedecor para análise de variância (ANOVA), em que consistiu em comparar a variação do preço do crude entre grupos e variação do preço de crude dentro de cada grupo, neste caso os grupos são os 12 meses.

As hipóteses do teste são:

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ (Ausência de sazonalidade)

H_1 : Pelo menos dois grupos têm médias diferentes

Estatística de teste é dada por:

$$F = \frac{\frac{SQE}{g-1}}{\frac{SQD}{n-g}} = \frac{QME}{QMD} \approx F_{(k-1, n-g)} \quad (27)$$

Onde:

$$SQE = \sum_{i=1}^g n_i (\bar{x}_i - \bar{x})^2 \quad (28)$$

$$SQD = \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x})^2 \quad (29)$$

Regra de decisão:

Seo valor obtido em (3.1) for inferior ao valor crítico de $F_{(g-1, n-g)}$, a um nível de significância de 5%, não há evidências suficientes para rejeitar a hipótese nula H_0 e conclui-se que há ausência de sazonalidade, isto é, o factor mês não influencia no comportamento das novas infecções pelo HIV, consequentemente, a razão de QME/QMD terá um valor próximo de 1. E também não há evidências suficientes para rejeitar H_0 se o p-value associado ao teste for maior que 5%. Para testar hipótese nula de igualdade das médias entre grupos através da estatística (3.1), deve-se cumprir os seguintes pressupostos.

3.2.2. Verificação das suposições

a) Normalidade:

O teste de Kolmogorov-Smirnov (K-S) é o teste estatístico adaptado para determinar, com uma certa margem de erro, se uma amostra é proveniente de uma população com distribuição contínua. A distribuição Normal é, pela sua importância nos modelos, aquela que com maior frequência pretendemos testar. A estatística do teste é,

$$DN = \max \left[\left| F(x_i) - F_0(x_i) \right|, \left| F(x_i) - F_0(x_i) \right| \right] \quad (30)$$

Onde:

A distribuição teórica acumulada sob H_0 é representada por $F_0(x_i)$ e a distribuição de frequências dos valores amostrais por $F(x_i)$. Como H_0 supõe que a amostra tenha sido obtida da distribuição $F_0(x_i)$ é razoável esperar que, para cada valor de x , $F(x_i)$ esteja próximo de $F_0(x_i)$, isto é, sob H_0 , espera-se que as diferenças entre $F(x_i)$ e $F_0(x_i)$ sejam pequenas.

Neste estudo, foi usado o teste de normalidade de K-S para testar o modelo normal com estimação de parâmetros foram corrigidos por Lillifors e é essa correcção que aparece na determinação do p-value (Sig.). O que se pretende testar é descrito pelas seguintes hipóteses:

H_0 : A população tem distribuição normal;

H_1 : A população não tem distribuição normal.

b) Homocedasticidade

Alguns procedimentos estatísticos comuns pressupõem que as variâncias das populações a partir das quais são desenhadas em diferentes amostras são iguais. Neste estudo foi usado o teste de Levene para avaliar este pressuposto. A estatística de teste de Levene (LM) dada a uma variável y com uma amostra do tamanho n dividido em k subgrupos, é dada por:

$$W = \frac{(n-k)}{(k-1)} \frac{\sum_{i=1}^k n_i (\bar{z}_i - \bar{z})^2}{\sum_{i=1}^k \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_i)^2} \approx F_{k-1, n-k} \quad (31)$$

Onde: n_i é o tamanho da amostra do i -ésimo subgrupo.

$$z_{ij} = |y_{ij} - \bar{y}_i| \quad (32)$$

$$\bar{z}_i = \frac{\sum_{j=1}^{n_i} z_{ij}}{n_i}, \quad \bar{z} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} z_{ij}}{n} \quad (33)$$

Testa se a hipótese nula de que as variâncias da população são iguais como descrito abaixo.

H_0 : As variâncias da população são iguais;

H_1 : As variâncias da população não são iguais.

Regra de desisão:

Conclui-se que há homogeneidade entre os grupos, se a um nível de significância de 5% não houver evidências suficientes para rejeitar a hipótese de que há igualdade de variâncias entre os grupos ou seja se o p-value for maior que nível de significância ou ainda se $W < F_{k-1, n-k}$. E se p-value for menor ou igual a 5%, indica que há evidências suficientes para rejeitar a hipótese de igualdade de variâncias entre os grupos.

3.2.3. Teste de estacionaridade

Segundo Gujarati (2000), o trabalho empírico baseado em dados de série temporal supõe que a série temporal envolvida seja estacionária. Mas em geral as séries temporais que envolvem dados de variáveis económicas não são estacionárias em níveis e para que sejam estacionárias devem ser integradas ou diferenciadas de modo a tornarem-se estacionárias ou seja se uma série temporal for integrada de ordem d , isto é, $I(d)$ suas d diferenças serão $I(0)$ ou seja estacionária.

Segundo Gujarati (2006), um procedimento concreto para a aplicabilidade de estatística de Dickey-Fuller, envolve estudar a natureza do processo de raiz unitária, por via dos seguintes processos aleatórios:

Y_t é um passeio aleatório:

$$\Delta y_t = \delta y_{t-1} + u_t \quad (34)$$

Y_t é um passeio aleatório com deslocamento:

$$\Delta y_t = \beta_1 + \delta y_{t-1} + u_t \quad (35)$$

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + u_t \quad (36)$$

Y_t é um passeio aleatório com deslocamento em torno de uma tendência estocástica.

3.2.4. Metodologia de Box-Jenkins

A metodologia Box & Jenkins é uma ferramenta importante de fácil aplicação para previsão de variáveis baseadas em séries temporais. A previsão do comportamento futuro dessas variáveis é de fundamental importância e ela ainda é pouco aplicada. Os modelos de Box e Jenkins partem da ideia de que, cada valor da série pode ser explicado por valores predeterminados. Estes modelos apresentam menor erro e consideram a série histórica, em ordem cronológica, em que se realizam as análises de autocorrelações e autocorrelações parciais, para se calcular uma estimação dos parâmetros, minimizando o erro quadrático.

Estes modelos possuem uma grande vantagem, que consiste em obter o erro quadrático mínimo apresentado pelo diagnóstico dos parâmetros estimados. A modelagem ARIMA é composta com as seguintes etapas:

1. Identificação do modelo

A identificação particular de um modelo ARIMA a ser ajustado aos dados pode ser considerado uma das fases mais críticas ao se utilizar uma modelagem ARIMA. A escolha do modelo a ser utilizado é feita principalmente com base nas funções de autocorrelações (FAC) e das autocorrelações parciais estimadas (FACP), que são utilizadas para comparar com as quantidades teóricas e identificar um possível modelo para os dados.

Esta fase é responsável por decidir se a série temporal é auto-regressiva de médias móveis ou mista. Esta etapa, é feita geralmente de forma visual, inspeccionando diagramas específicos obtidos nos dados, ou empregue nas várias técnicas estatísticas. Essa é considerada a etapa mais difícil e delicada visto que não há consenso sobre qual é a melhor estratégia a ser seguida. Entre as várias existentes, duas se destacam: A análise das funções de autocorrelação e autocorrelação parcial amostrais e o uso de um critério de selecção de modelos (Critérios de Akaike ou Bayesiano).

Tabela 1. Propriedades das FAC e FACP nos modelos ARIMA

Modelos	FAC	FACP
(1, d, 0)	Decaimento exponencial ou oscilatório	$\epsilon_{kk} = 0$, para $k > 1$
(2, d, 0)	Decaimento exponencial ou senoidal	$\epsilon_{kk} = 0$, para $k > 2$
(p, d, 0)	Decaimento exponencial e/ou senoidal	$\epsilon_{kk} = 0$, para $k > p$
(0, d, 1)	$\rho_k = 0$, para $k = 1$	Dominado por decaimento exponencial
(0, d, 2)	$\rho_k = 0$, para $k = 2$	Dominado por decaimento exponencial ou senoidal
(0, d, q)	$\rho_k = 0$, para $k = q$	Dominado pela combinação linear do decaimento exponencial ou senoidal
(1, d, 1)	Decaimento exponencial a partir da defasagem 1	Dominado por decaimento exponencial a partir da defasagem 1
(p, d, q)	Decaimento exponencial ou senoidal depois da defasagem p-q	Dominado por decaimento exponencial ou senoidal depois da defasagem p-q

Fonte: Bruno (2001)

Para modelos sazonais, o comportamento da FAC e da FACP deve ser analisado também próximo da defasagem sazonal, por exemplo, defasagem 12 para dados mensais e defasagem 4 para dados trimestrais. A Tabela abaixo resume as propriedades da FAC e da FACP para modelos SARIMA.

Tabela 2- Propriedades das FAC e FACP nos modelos ARIMA

Modelos	FAC	FACP
(P, D, 0)	Decaimento	Picos nas defasagens s, 2s, ..., Ps e corte após Ps
(2, D, 0)	Picos nas defasagens s, 2s, ..., Qs e corte após Qs	Decaimento
(P, D, 0)	Picos nas defasagens s, 2s, ..., Qs e corte após Qs	Picos nas defasagens s, 2s, ..., Ps e corte após Ps
(0, D, 1)	Decaimento rápido na defasagem sazonal	Decaimento rápido na defasagem sazonal
(0, D, 2)	Valores pequenos em todas as defasagens sazonais (Não há picos)	Valores pequenos em todas as defasagens sazonais (Não há picos)

Fonte: Bruno (2001)

2. Estimação dos parâmetros

Identificada a ordem de um modelo ARIMA para uma série temporal precisa-se agora estimar os parâmetros deste modelo. Para isso, pode-se utilizar método dos momentos, estimadores de mínimos quadrados e estimadores de máxima verossimilhança.

3. Diagnóstico ou verificação do modelo

Essa etapa consiste em verificar se o modelo identificado e estimado é adequado. Em caso positivo, ele pode ser utilizado para fazer previsões. Em caso negativo, será necessário identificar outro modelo e repetir as etapas de estimativa e verificação. Segundo Fava (2000), as formas de verificação frequentemente utilizadas são: análise de resíduos e avaliação da ordem do modelo. Alguns testes de adequação do modelo são:

- Teste de autocorrelação residual;
- Teste de Box-Pierce;
- Teste da autocorrelação cruzada;
- Análise de resíduos.

A avaliação feita por meio dos resíduos do modelo estimado, exige que os mesmos devem apresentar estimativas do ruído branco. O teste de ruído branco é feito verificando o comportamento da sua FAC para os períodos estimados. Essa conclusão é reforçada após a utilização do teste de Ljung-Box Q^* , que não rejeita, para qualquer defasagem k , a hipótese de erros não correlacionados. Assim, podemos considerar que os resíduos se comportam como ruído branco.

4. Avaliação da ordem do modelo

Para além disso, a análise do desvio padrão residual e outro indicador para verificar se a ordem do modelo é adequado, quanto menor for o erro padrão, melhor serão as previsões.

Segundo Costa (2005), caso o modelo não seja adequado, o ciclo é repetido, voltando a fase da identificação. Contudo, o procedimento muito utilizado é identificação de vários modelos, mas alguns deles são dados estimados e, escolhe-se um. Quando se obtém um modelo satisfatório, passa-se para a última etapa da metodologia: realizar previsões.

O objectivo desta avaliação é verificar se um modelo não apresenta parâmetros em excesso, esta observação é feita com base no erro padrão. Se o valor do coeficiente for pequeno em relação a seu erro padrão, conclui-se que ele não é significativo.

5.Previsão apartir do modelo

Depois da identificação e estimação de um modelo ARIMA adequado às observações, procede-se com avaliação dos métodos que possam ser utilizados na modelagem ARIMA para prever os valores das observações para os próximos meses. Esta etapa não requer grande intervenção por parte analista, uma vez identificado o modelo, as estimativas dos respectivos parâmetros são feitas automaticamente recorrendo a um dos métodos: máxima verosimilhança, mínimos quadrados e momentos.

Importa ressaltar que previsões utilizando modelos ARIMA são eficazes para um período curto e as melhores previsões serão aquelas que apresentam um erro quadrático médio (EQM) mínimo. Além dessa medida o desempenho preditivo dos modelos foi avaliado segundo o tamanho do erro de previsão, mensurado pela raiz quadrada da média dos quadrados do erro (RMSE) e (MAPE).

- **Raiz quadrada da média dos quadrados do erro (RMSE)**

$$RMSE = \sqrt{\frac{\sum_{i=1}^n e_t^2}{n}} \quad (37)$$

- **Média do erro absoluto (MAE)**

$$MAE = \frac{\sum_{i=1}^n |e_t|}{n} \quad (38)$$

- **Erro absoluto percentual médio(MAPE)**

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{\mu_t}{y_t} \right|}{n} \quad (39)$$

Onde:

$$e_t = y_t - \hat{y}_t$$

Essas duas medidas foram usadas para certificar se do modelo que tem a melhor capacidade para fazer a previsão. Onde as medidas que podem ser usadas para verificar quais modelos que têm as melhores medidas RMSE e erro absoluto percentual médio MAPE (este não vale para dados próximos de zero, sendo preferível a utilização de outro método para a análise dos erros). Os valores do erro percentual absoluto médio podem ser relacionados com a potencialidade das previsões: MAPE <10\% - previsão é potencialmente muito boa, MAPE <20\% - previsão é potencialmente boa, MAPE <30\% - previsão é potencialmente razoável e MAPE >30\% - previsão é potencialmente inexata.

Capítulo IV

Resultados e Discussão

Neste capítulo, fez-se a apresentação e a discussão dos principais resultados obtidos com a aplicação dos modelos descritos na secção metodológica. Inicialmente faz-se uma análise descritiva da série dos casos de infecção por HIV/SIDA, o comportamento da série através de uma análise gráfica, teste de ADF para a estacionariedade da série, comparação entre os melhores modelos e selecção do melhor com o propósito de realizar a previsão.

4.1 Resultados

4.1.1. Análise descritiva

A Tabela (3), mostra o comportamento da série dos casos positivos no teste de HIV/SIDA no período em análise, onde observa-se uma média mensal de 3554 casos positivos com desvio padrão aproximadamente de 1347 casos positivos no mesmo período. O número máximo de casos de infecção por HIV por mês, foi de 8240 e o número mínimo de casos foi de 693 em um mês.

Tabela 3. Estatísticas descritivas da série das infecções

Variável	Observações	Média	Desviu padrão	Coeficiente de variação	Observação mínima	Observação máxima
Infecções	N = 77	3554	1347	37.9%	693	8240

Analisando o coeficiente de variação, verifica-se que existe uma grande variabilidade nos casos de infecção por HIV/SIDA ao longo do período em análise, ou seja, apresentam uma dispersão alta, pois o coeficiente de variação é igual à 37.9% sendo que é maior que 30% que corresponde a dispersão alta. Examinando a série temporal de casos de infecção por HIV/SIDA na Província da Zambézia na Figura (4), pode-se afirmar que o gráfico apresenta comportamentos diferenciados durante o tempo, visto que em 2022 registam-se altos casos notificados de infecção de HIV/SIDA, situação contraditória ao período inicial de 2018.

A mesma série não apresenta tendência crescente ou decrescente, como ilustra o gráfico abaixo, e pode-se suspeitar que ela se desenvolve de forma aleatória ao longo do tempo, sugerindo que ela seja estacionária, mas para efeitos de confirmação foi feito o teste de raiz unitária.

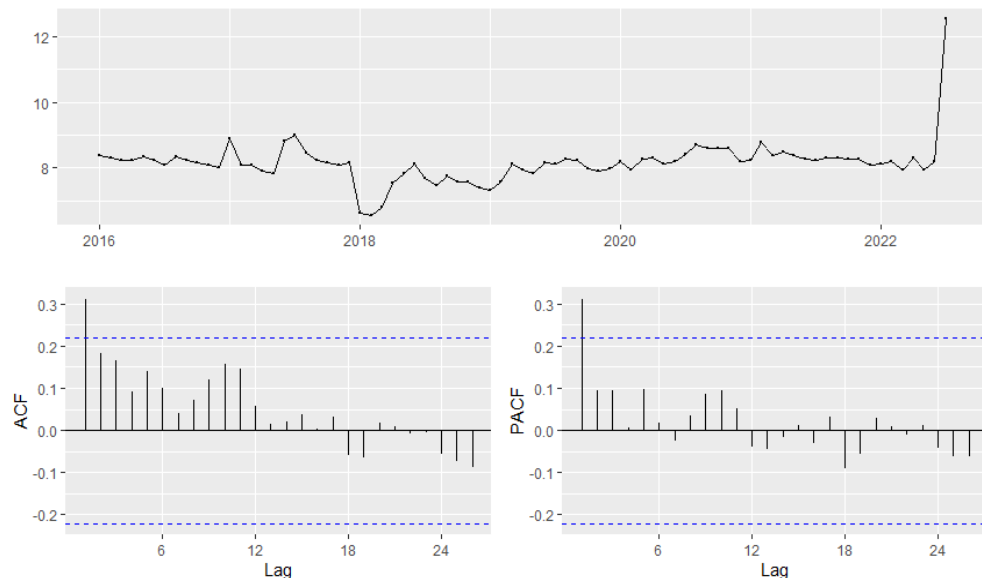


Figura 2 Comportamento da série de casos de infecção por HIV/SIDA

Entretanto, antes mesmo de fazer todas as análises da série em estudo, incluindo o teste de de raiz unitária que é o primeiro teste a realizar antes de fazer qualquer outro teste, os dados foram submetidos a verificação de algumas suposições.

4.1.2. Verificação das suposições

Antes de proceder com as análises é importante verificar o pressuposto da normalidade, visto que é a fase fundamental para averiguar a qualidade do ajustamento da distribuição normal.

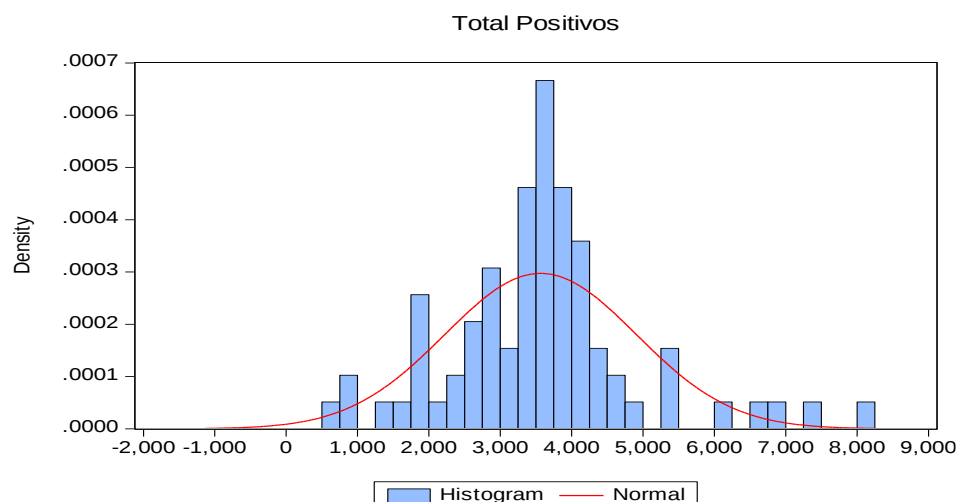


Figura 3 - Histograma da série de casos de infecção por HIV/SIDA

A representação gráfica dos dados é geralmente o primeiro passo da análise de toda série temporal. Analisando a Figura 3, verifica-se que os valores dos coeficientes de assimetria e de curtose são ambos estatisticamente diferentes da distribuição normal (0 e 3 respectivamente). Ainda na figura 4, também é possível visualizar o carácter leptocurtico da distribuição dos casos positivos de HIV.

Tabela 4. Teste de normalidade de Kolmogorov-Smirnov

Variável	Estatística K-S	Df.	Sig.
Casos positivos	0.144	78	0.546

O teste não paramétrico de Kolmogorov-Smirnov, Tabela 4, mostra que a série de casos de infecção por HIV/SIDA na Província da Zambézia apresenta o valor do teste e a probabilidade associada de 0.546. Assim, há evidência suficiente para se Rejeitar a hipótese de que a distribuição é normal, ou seja, a série dos casos de infecção por HIV/SIDA na Província da Zambézia entre Janeiro de 2016 e Junho de 2022 possui distribuição normal.

O teste de igualdade das médias dos casos de infecção por HIV/SIDA, sugere a não rejeição da hipótese nula de igualdade das médias mensais de casos positivos de HIV na Província de Zambézia, o que significa que a um nível de significância de 5% não existem diferenças significativas entre as médias mensais dos casos positivos de infecção por HIV/SIDA registados na Província de Zambézia no período em análise, visto que o valor do sig (0.879) é maior do que $\alpha = 0.05$ (nível de significância), como descrito na Tabela 5.

Tabela 5. Anova para comparação das médias dos casos positivos de HIV

Regressão	Soma dos quadrados	Df.	Quadrático Médio	Estatística F	Sig.
Entre grupos	11175799,368	11	1015981,761	0.526	0.879
Dentro dos grupos	127524916,786	66	1932195,709	-----	-----
Totais	1387007616,154	77	-----	-----	-----

Mediante o pressuposto da aplicação do teste de Levene, a Tabela 6, apresenta a não rejeição da hipótese nula da homogeneidade na variância dos casos de infecção por HIV/SIDA na Província da Zambézia, ou seja, as diferenças de homogeneidade das variâncias observadas nos casos de infecção por HIV/SIDA na Província da Zambézia, não são consideradas estatisticamente significativas visto que o p-value é maior que o nível de significância aplicado neste trabalho 5%.

Tabela 6. Teste de Homogeneidade das variâncias

Estatística de Levene	Df1	Df2	Sig.
Casos positivos	11	66	0.654

Desta feita, fez-se a análise da estacionariedade da série em análise, sob o risco de proceder-se com a análise e obter-se uma regressão sem sentido, o que levaria a uma previsão inadequada se futuros casos de infecções pelo HIV/SIDA.

4.1.3. Análise da Estacionariedade

A análise da função de autocorrelação (FAC) e da função de autocorrelação parcial (FACP) descritas na figura (), não permite tirar conclusões sobre a estacionariedade da série de casos de infecção por HIV/SIDA na Província da Zambézia no período de Janeiro de 2016 à Junho de 2022, sendo assim realizou-se o teste de ADF para análise da estacionariedade.

Tabela 7. Teste de ADF das infecções registadas em nível

Série	Estatística t	$\tau (\alpha = 0.05)$	lags	Sig.
Casos positivos	-1.908	-2.89	74	0.169

O valor da estatística t é de -1,8102, que é menor do que o valor crítico à significância de 5%, que é de -3,43 quando comparados em termos absolutos. Portanto, com base na análise gráfica, do correlograma ACF e no p-valor associado ao teste de Dickey-Fuller, a conclusão é de que, para os períodos janeiro de 2016 à Junho de 2022, a série em nível dos casos positivos foi não estacionária; O que torna indispensável a realização das 1ªs diferenças, de modo a torna lá estacionária.

Tabela 8. Teste de ADF das infecções registadas em nível

Série	Estatística t	$\tau (\alpha = 0.05)$	lags	Sig.
Casos positivos	-4.431	-2.89	73	1.21e-05

Depois das 1ªs diferenças pode-se verificar que estatística t -4.431 é superior que o $\tau = -2.89$ em termos absolutos e o p-valor também é inferior que o nível de significância e pode-se concluir que a em análise é estacionária nas suas 1ªs diferenças ou então, pode-se afirmar que a série é I(1).

4.1.4. Identificação do modelo

Neste estudo, as sugestões dos picos, decaimentos e outras que foram descritos como indispensáveis para a identificação do modelo foram ignoradas, pois, o pacote estatístico R, dispõe de um método ou função denominada **auto.arima()**, que consiste em fazer diversos modelos com o auxílio computacional e trazer todas as combinações de modelos com os seus respectivos parâmetros, cabendo apenas ao investigador avaliar aquele que apresentar o menor AIC entre os modelos gerados.

Deste modo, passa-se a apresentar os 11 modelos gerados pela função **auto.arima**, de salientar que a maiorias dos modelos contém a contante, com excepção de apenas um modelo, além disso, foram apresentados os seus respectivos AIC para efeitos de avaliação do melhor modelo, como é demonstrado na seguinte tabela:

Tabela 9. Modelos gerados pela função **auto.arima()**

Modelo ajustado	Constante	AIC
SARIMA(2,1,2)(1,0,1)[12]	Com constante	Inf
ARIMA(0,1,0)	Com constante	144.0246
ARIMA(1,1,0)(1,0,0)[12]	Com constante	148.0174
ARIMA(0,1,1)(0,0,1)[12]	Com constante	148.0347
ARIMA(0,1,0)	Sem constante	142.5400*
ARIMA(0,1,0)(1,0,0)[12]	Com constante	145.7951
ARIMA(0,1,0)(0,0,1)[12]	Com constante	145.8150
ARIMA(0,1,0)(1,0,1)[12]	Com constante	148.0016
ARIMA(1,1,0)	Com constante	146.1889
ARIMA(0,1,1)	Com constante	146.1889
ARIMA(1,1,1)	Com constante	Inf

De entre os modelos apresentados, o melhor modelo é o ARIMA (0,1,0), visto que, esse modelo possui um menor valor de AIC, como ilustra a tabela (9). Assim sendo, prosseguiu se com a estimação do modelo que será usado para fazer as previsões, esse modelo também pode ser denominado como o modelo integrado de primeira ordem.

4.1.5. Diagnóstico do modelo estimado

Após estimar o modelo é necessário verificar se o mesmo é adequado. Se o modelo estiver adequadamente especificado, os resíduos do modelo estimado serão estimativas de ruído branco e assim, os coeficientes de Auto correlação dos resíduos devem estar estatisticamente iguais a zero.

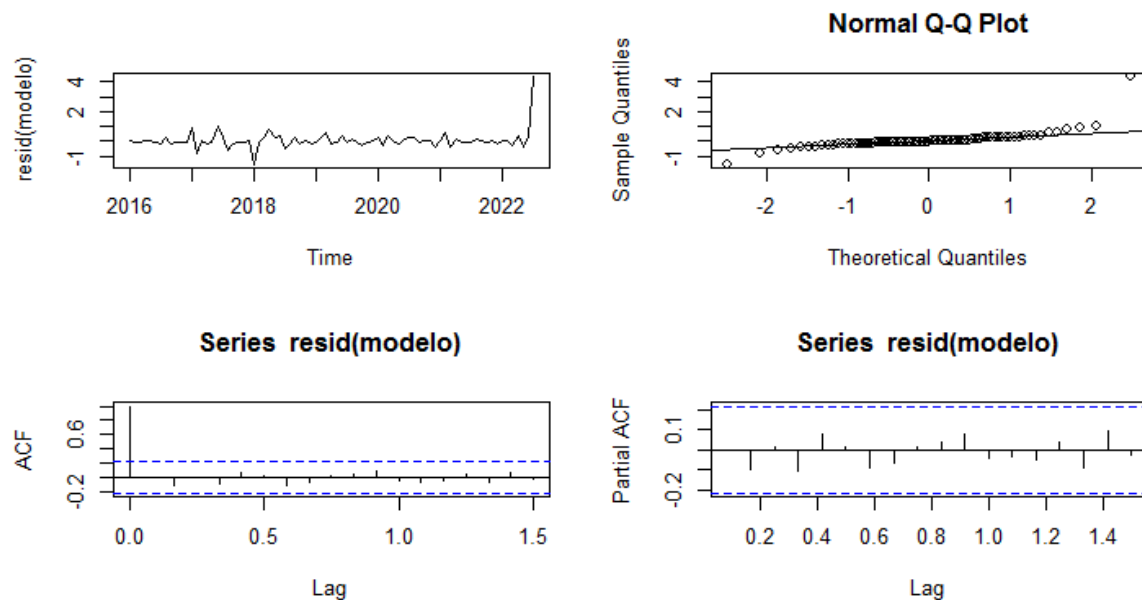


Figura 4 - Representação gráfica dos resíduos do modelo

Pode se verificar no primeiro gráfico descreve a série dos resíduos do modelo estimado, e no canto superior direito, foi representado o Q-Q plot que sugere a normalidade dos resíduos, pois, a vários pontos estão sobre a recta estimada, as duas representações abaixo da figura(8) ilustra os correlogramas da FAC e FACP, estas, apresentam as defasagens dentro dos limites, levando a suposição de que os resíduos apresentam um ruído branco. Para provar esta suposição foi preciso realizar os testes formais dos resíduos do modelo estimado e verificou se o seguinte:

Tabela 10. Teste de ADF dos resíduos do modelo estimado

Série	Estatística t	$\tau (\alpha = 0.05)$	lags	Sig.
Casos positivos	-4.539	-2.89	73	1.45e-05

Para verificar as autocorrelações parciais dos resíduos do modelo estimado, foi realizados dois testes que são o teste de Box-Piece e Ljung-Box, desta feita passa se a apresentar os resultados nas tabelas que se seguem.

Tabela 11. Teste de Box-Piece dos resíduos do modelo

Estatística Q_{bp}	Lags usados	Sig.
3.6534e-06	1	0.9985
0.80489	2	0.6687

Tanto a tabela (11) e a tabela (12) pode se verificar que ambos os testes, apresentam os p-valores associados aos testes superiores ao nível de significância ($\alpha=0.05$), então não rejeitam a hipótese nula de que as correlações parciais até o número de lags estudados são iguais a zero, deste modo, pode se dizer que apartir dos testes Box-Piece e Ljung-Box não existe a multicolinearidade entre os diferentes registos de casos positivos do HIV/SIDA.

Tabela 12. Teste de Ljung-Box dos resíduos do modelo

Estatística Q_{lp}	Lags usados	Sig.
3.7957e-06	1	0.9984
0.84726	2	0.6547

Depois de ter se feito o diagnóstico e verificado que o modelo estimado é adequado passa-se para a última etapa da metodologia box-Jenkins, que consiste em fazer as previsões através do melhor modelo ajustado.

Para poder avaliar o desempenho do modelo foram escolhidos alguns modelos incluindo o escolhido e foram apresentadas algumas medidas dos erros estimados pelos modelos. Deste modo na tabela (13), pode-se verificar que os valor do MAPE 0.2186287, o que quer dizer que, as previsões feitas pelo modelo em análise (ARIMA(0,1,0)), são considerados potencialmente razoáveis, pois, o valor de MAPE é inferior a 30% (0.3).

Tabela 13. Medidas para avaliar a capacidade de prever do modelo estimado

Modelo	BIC	RMSE	MAPE
ARIMA(0,1,0)	144.84	0.5917052	0.2186287
ARIMA(2,1,2)(1,0,1)[12]	162.98	0.5508111	3.358672
ARIMA(1,1,0)(1,0,0)[12]	153.16	0.5896649	0.5896649

Deste modo, antes de fazer a previsão, procedeu-se com a verificação dos pressupostos que devem ser cumpridos na metodologia Box & Jenkins, e verificou se que todos os testes sobre os resíduos de maneira a certificar-se da qualidade do modelo ajustado, assim as especificações são adequadas e já pode se proceder com as previsões.

4.1.6. Previsão

Uma vez já escolhido o melhor modelo dentre os sugeridos, passa-se a fase final do processo que é a previsão de casos de infeção por HIV/SIDA na Província da Zambézia (julho 2022 ao junho de 2023).

Graficamente são ilustradas as Infecções previstas e seu intervalo de confiança na área pintada a azul na figura (5).

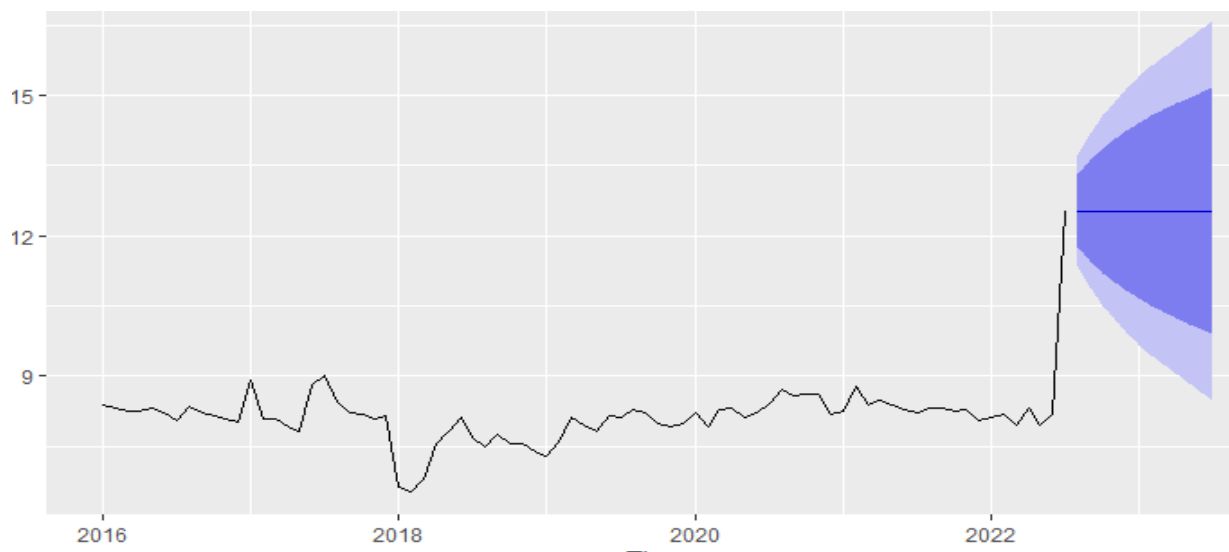


Figura 5 – Representação grafica das previsões num intervalo de confiança de 95%.

Na Tabela (5), são apresentadas as previsões dos casos de infecção por HIV/SIDA na Província da Zambézia para os próximos 12 meses em relação ao período em análise. De salientar que os valores mensais previstos são o acumulo dos meses anteriores por isso, a tendência é crescente.

Tabela 14. Previsão dos casos de infecção por HIV/SIDA

Período	Limite Inferior	Previsão das infecções	Limite Superior
Julho/ 2022	11.368454	13.29873	13.70272
Agosto/ 2022	10.885012	13.61484	14.18616
Setembro/ 2022	10.514054	13.61484	14.55712
Outubro/ 2022	10.514054	13.85739	14.55712
Novembro/ 2022	10.201322	14.06188	14.86985
Dezembro/ 2022	9.925800	14.24203	15.14537
Janeiro/2023	9.676709	14.40490	15.14537
Fevereiro/2023	9.447646	14.55468	15.62352
Março/2023	9.234439	14.69409	15.83673
Abril/2023	9.034191	14.82502	16.03698
Maio/2023	8.844791	14.94887	16.22638
Junho/2023	8.664648	15.06665	16.40652

Capítulo V

Conclusões e Recomendações

Neste capítulo, foram apresentadas as principais conclusões e recomendações consideradas pertinentes depois dos resultados obtidos apartir do estudo realizado, onde foram analisados os números de infectados pelo HIV/SIDA em Zambézia, esses dados foram colhidos mensalmente no período de janeiro de 2016 à junho de 2022.

5.1. Conclusões

Depois de analisados os dados registados das infecções mensais desde 2016 à 2022, onde foi usando a metodologia de Box-Jenkins, chegou-se as seguintes conclusões:

- A série de casos de infecção por HIV/SIDA na Província da Zambézia entre Janeiro de 2016 e Junho de 2022 é estacionária em nível apesar de apresentar flutuações em alguns meses, por apresentar dados dispostos de forma exponencial, fez se uma tranformação logartmica da série das infecções registadas mensalmente na província da Zambézia;
- Foram analisados cerca de 11 modelos, tendo o modelo **ARIMA(0,1,0)**, demonstrado evidências de apresentar um melhor desempenho estatístico, pois, este apresenta os melhores indicadores, tanto do $BIC = 144.84$, $RMSE = 0.5917052$ e $MAPE = 0.2186287$, estes resultados fazem desse modelo o melhor quando comparado aos outros modelos.
- resultados no que diz respeito as previsões fora da amostra, sendo portanto, o modelo escolhido para representar a série de casos de infecção por HIV/SIDA na Província da Zambézia entre Julho de 2022 e Junho de 2023, onde verificou se que todos os valores previtos encontram se dentro dos limites do intervalo de confiança de 95%, o que inspira mais confiabilidade dos valores previstos;
- As previsões no período de Julho de 2016 à Junho de 2023, demostram uma tendência crescente de novas infecções pelo HIV/SIDA na Província da Zambézia.

5.2. Recomendações

Recomenda-se para estudos futuros neste campo, sejam tomadas todas as províncias do País e que os resultados sejam comparados com a realidade por forma a fortalecer a escolha do modelo adequado para a série, o que não foi feito neste trabalho devido a não existência da informação referente ao período de previsão;

Recomenda-se mais ainda, a utilização de outros métodos para fazer previsão e comparados aos obtidos com a metodologia de Box-Jenkins.

Referências Bibliográficas

1. Alencar, L. (2010). Estudo sobre a forma de ajuste do índice de Basileia pelos bancos no Brasil. Relatório de Economia Bancária e Crédito 2010, Banco Central do Brasil, Disponível em: <http://www.bcb.gov.br/Pec/Depep/Spread/REBC2010.pdf>;
2. Almeida, D. S. O.; Costa, R. L.; Silva, T. M. (2005). Chega de tabu! A sexualidade sem medos e sem cortes. Disponível em: <http://www.unesp.br/prograd/PDFNE2005/artigos/capitulo%201/ch>. Acesso em: 12 Setembro 2023;
3. Barros, M. (2005). Processo Estocástico. São Paulo: Publit;
4. Brito, A. M.; Castilho, E. A.; Szwarcwald, C. L. (2000). AIDS e infecção pelo HIV no Brasil: uma epidemia;
5. Brasil. Ministério da Saúde (2010). Departamento de DST, Aids e Hepatites Virais. Recomendações para terapia antirretroviral em adultos infectados pelo HIV. Brasília;
6. Brockwell, P.J. e R.A. Davis (1996). Introduction to Time Series and Forecasting. New York, Springer-Verlag;
7. Chatfield C. (2004). The analysis of time series: An introduction. 5. ed. Chapman and Hall/CRC;
8. CNCS / CRIS (2004). Implementação de Acções de Combate ao HIV/SIDA em Moçambique, entre Janeiro a Maio de 2004;
9. Costa, A. (2005). Mathematical/Statistical and Physical/Meteorological Models for Short-term Prediction of Wind Farms Output. Tese (Doutorado em Engenharia) - Escuela Técnica Superior de Ingenieros Industriales, Univesidad de Madrid, Madrid;
10. Enders, W. (2004). Applied Econometric Time Series. Iowa State University. 2ª ed;
11. Franco, G. H. B. (2005) “Auge e declínio do inflacionismo no Brasil”. In: Giambiagi, F. et al (Org.). Economia brasileira contemporânea (1945-2004), 1ª edição. Rio de Janeiro: Campus Elsevier;
12. Granger, C. W. J.; Newbold, P. (1986). Forecasting Economic Time Series. London, Academic Press, 211p;
13. Hill, R. C., Griffiths, W. E., Judge, G. G. e de Farias, A. A. (2010). Econometria, Saraiva;

14. Johnston, J. & Dinardo, J. (2001), Métodos econométricos. tradução: Manuela magalhães hill, fátima ferrão, rui menezes
15. UNAIDS (2015). Country Profile 2015. Disponível em: http://www.unaids.org/en/Regions_Countries/Countries/mozambique.asp. Acesso em 28 ago. 2015.
16. Governo de Moçambique (2015). Plano Estratégico Nacional de Resposta ao HIV e SIDA 2015 – 2019. Maputo, Moçambique;
17. Gujarati, D. N. (2006) . Econometria Básica. 4ª Edição, editora Campus;
18. Liebel, M. J. (2004). Previsão de receitas tributárias: o caso do ICMS no estado do Paraná. Dissertação (Mestrado Profissionalizante em Engenharia) - Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre;
19. Luz, M.L.S.; Dalpasquale, V.D.; Scapim, C.A.; Lucca-Braccini, A.; Royer, M.R.; Mora, F. (2005). Influência da umidade das sementes na capacidade de expansão de três genótipos de milho-pipoca (Zeamays L.). Acta Scientiarum Agronomy, v.27, n 3, p.549- 553;
20. Morettin, P.A.; Toloi, C.M.C. (2004). Análise de Séries Temporais. São Paulo: Editora Edgard Blucher Ltda;
21. Sousa, R. F. (2006). Impacto do Preço do Petróleo na Política Energética mundial. Dissertação apresentada para obtenção do grau de mestre. Disponível em <http://www.ppe.ufrj.br/ppes/production/tesis/souzafr.pdf>. Acedida a 25/07/2023;
22. Pellegrini, F. R. E Fogliatto, F. S. (2001). Passos para implantação de sistemas de previsão de demanda – Técnicas e estudo de Casos. Revista Produção. 11(1), pp.43-64;
23. Werner, L e Ribeiro, J. L. D. (2003). Previsão de demanda: uma aplicação dos modelos box-jenkins na área de assistência técnica de computadores pessoais Revista Gestão e Produção. 10(1), pp.48-67;

Anexos

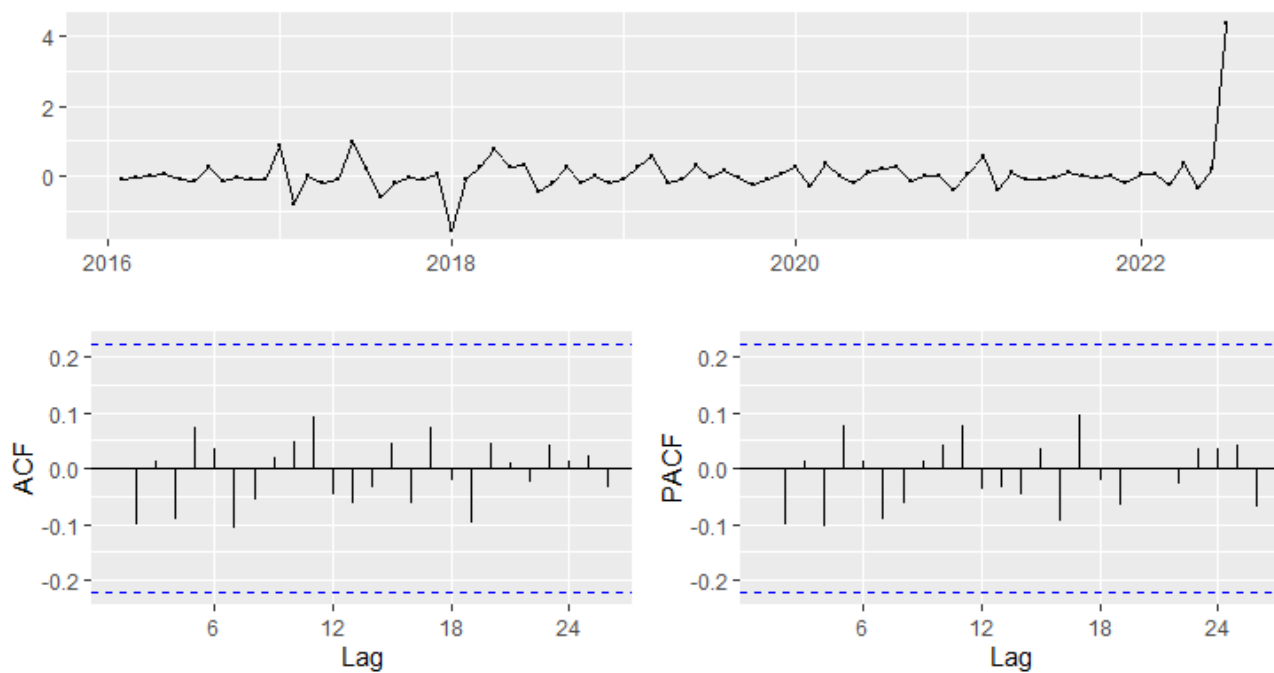


Figura 6 - Comportamento das 1^{as} diferenças da série de casos de infecção por HIV/SIDA

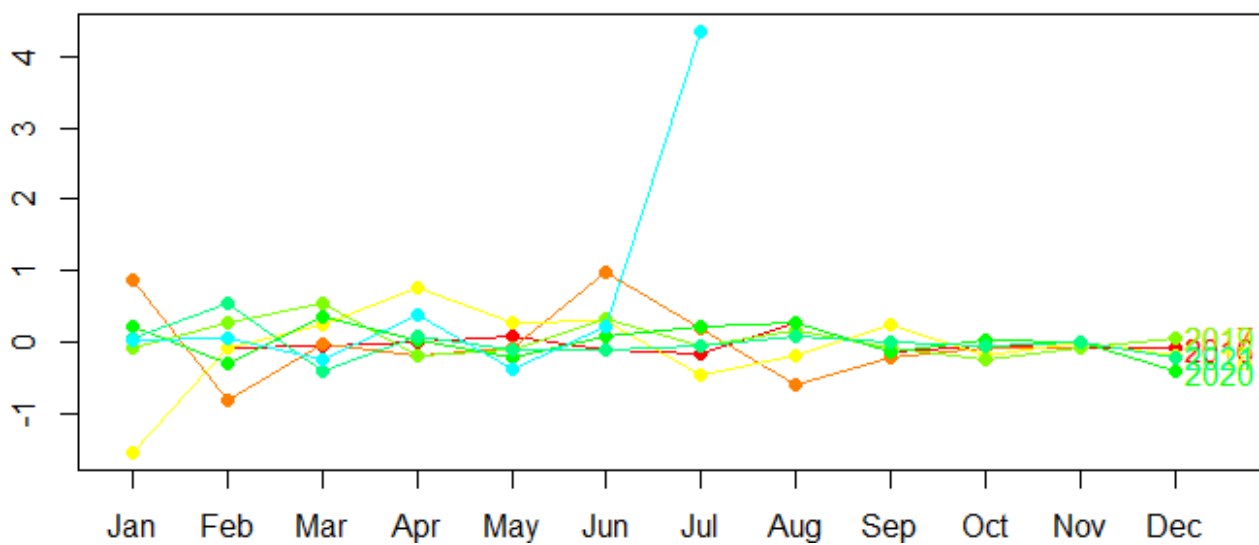


Figura 7 - Comportamento da sazonalidade das 1^{as} diferenças de série de casos de infecção por HIV/SIDA

