



UNIVERSIDADE EDUARDO MONDLANE

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE MATEMÁTICA E INFORMÁTICA

CURSO DE ESTATÍSTICA

Trabalho de Licenciatura

TEMA

**MODELAGEM ARIMA: UMA APLICAÇÃO NA ANÁLISE E PREVISÃO DO VOLUME
DE CLIENTES DO SERVIÇO TELEFÓNICO PÓS-PAGO PROFISSIONAIS DAS TDM,
SARL, NA PROVÍNCIA DE MAPUTO**

Autora: Salmina Leonardo Moiane

Supervisor: Dr. Milagre Mula

Maputo, Novembro de 2008

Modelagem ARIMA: Uma aplicação na análise e previsão do volume de clientes do serviço telefónico pós-pago profissionais das TDM, SARL, na província de Maputo

Salmina Leonardo Moiane

TRABALHO DE LICENCIATURA EM ESTATÍSTICA

MODELAGEM ARIMA: UMA APLICAÇÃO NA ANÁLISE E PREVISÃO DO VOLUME DE CLIENTES DO SERVIÇO TELEFÓNICO PÓS-PAGO PROFISSIONAIS DAS TDM, SARL, NA PROVÍNCIA DE MAPUTO

Trabalho apresentado ao curso de Estatística do Departamento de Matemática e Informática da Faculdade de Ciências da Universidade Eduardo Mondlane como requisito para a obtenção do título de Licenciada em Estatística.

Supervisor:

Dr. Milagre Mula

Maputo

Novembro de 2008

Salmina Leonardo Moiane

DECLARAÇÃO DE HONRA

Declaro que o presente trabalho é resultado da minha própria investigação, e que não foi submetido para outro grau que não seja o indicado - Licenciatura em Estatística na Universidade Eduardo Mondlane.

Maputo, Novembro de 2008.

Salmina Leonardo Moiane
(Salmina Leonardo Moiane)

DEDICATÓRIA

Ao meu marido José Afonso Mendes, pelo afecto, paciência, encorajamento e pelo auxílio constante prestado.

Aos meus filhos Hamilton, Elton e José Hélder, maiores incentivadores, por todo carinho e pela paciência e compreensão que tiveram em todas as horas de ausência.

AGRADECIMENTOS

A DEUS;

Ao meu pai Johane Moiane Leonardo, agradeço por este trabalho, como minha mais saudosa homenagem;

A minha mãe Laurinda Boane, a família Mendes e a família Moiane, pelo carinho e apoio irrestrito na minha trajectória;

Ao meu supervisor Dr. Milagre Mula, aos docentes, pelas oportunidades, ensinamentos e ideias. Aos colegas de curso de estatística, que juntos estivemos nesta formação;

Por último agradeço a todos que directa ou indirectamente apoiaram-me.

Muito obrigada!

“O que sabemos é uma gota,
o que ignoramos é um oceano”

Isaac Newton
(1643-1727)

RESUMO

O objectivo deste estudo foi identificar um modelo de previsão para o volume de clientes do serviço telefónico pós-pago profissionais da Empresa Telecomunicações de Moçambique, SARL como ferramenta auxiliar ao planeamento de suas acções. Por ser o volume de clientes um dos factores que dita o desenvolvimento da empresa, é de interesse conhecer o comportamento futuro do volume de clientes em específico. Para isso, utilizou-se a Metodologia estatística para séries temporais, mais especificamente o método Box-Jenkins.

Os dados tratados foram fornecidos pela TDM, SARL, Direcção Comercial, e se referem aos valores mensais do número de clientes dos serviços telefónicos pós-pago profissionais, na província de Maputo, durante o período de 2001 a Agosto de 2008. Para realizar as análises, fez-se uso dos aplicativos computacionais Eviews e SPSS. Vários modelos de previsão foram avaliados, resultando um modelo de previsão.

Para se atingir o objectivo, inicialmente foi feita a análise da série temporal com a utilização da metodologia ARIMA – Modelos de Box-Jenkins, chegando-se a definição de modelos matemáticos que podem ser usados para previsão. De acordo com os resultados das análises, a série do volume de clientes pode ser representada pelo modelo ARIMA (7,1,0);

As previsões para o período de 05/2008 a 12/2008, usando o modelo ARIMA (7,1,0) não se diferenciam bastante dos valores observados do volume de clientes;

Palavras-chave: *Séries temporais, volume de clientes, Box-Jenkins, Previsão.*

LISTA DE FIGURAS

Figura 2.1 – ACF e PACF para um Modelo AR(1)	24
Figura 2.2 – ACF e PACF para um modelo MA (1).....	24
Figura 2-3 – Fluxograma do Modelo ARIMA de Box e Jenkins.....	26
<i>Figura 4-1. Evolução temporal do volume de clientes nos serviços telefônicos pós - pago profissionais das TDM, SARL em nível, 2001-1 a 2008-8.</i>	35
Admite-se que esta alteração estrutural na evolução temporal do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL se deva:	35
Figura 4-2: <i>Histograma do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8.</i>	36
Figura 4-3. <i>Evolução temporal do volume de clientes nos serviços telefônicos pós-pago profissionais da TDM, SARL em primeiras diferenças, 2001-1 a 2008-8.</i>	37
Figura 1A – <i>Gráfico de Auto correlação da Série do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8.(em nível)</i>	52
Figura 2A – <i>Gráfico de Auto correlação da Série do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8.(em primeiras diferenças)</i>	52
Figura 1C: <i>Gráfico dos resíduos do modelo AR (7)</i>	57
Figura 2C: <i>Correlograma da FAC e FACP dos resíduos do modelo AR (7)</i>	57
Figura 3C: <i>Gráfico dos resíduos do modelo AR (7)</i>	58
Figura 4C: <i>Correlograma da FAC e FACP dos resíduos do modelo AR(7)</i>	58
Figura 5C: <i>Gráfico dos resíduos do modelo AR (7)</i>	59
Fonte: Adaptação da autora	59
Figura 6C: <i>Correlograma da FAC e FACP dos resíduos do modelo AR (7)</i>	59

LISTA DE TABELAS

Tabela 3-1: . Padrões teóricos das funções de autocorrelação e auto-correlação parcial.....	32
Tabela 4-1. <i>Teste de raiz unitária em nível com intercepto e tendência para o logaritmo do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8.</i>	38
Tabela 4-2. <i>Teste de raiz unitária em primeiras diferenças com intercepto e tendência para volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8.</i>	38
Tabela 4-3: <i>Estatísticas de estimação de modelos</i>	43
Tabela 4-4. <i>Previsão do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL segundo o processo AR(7) para o período de Maio a Dezembro de 2008.</i>	44
Tabela 4-5: <i>Valores observados e previstos do volume de clientes no período de Maio a Agosto/08</i> ..	45
Tabela 1A: <i>Volume de clientes dos serviços telefônicos pós-pago profissionais de 01/2001 a 08/2008</i>	51
Tabela 2A: <i>Estatísticas descritivas do volume de clientes dos serviços telefônicos pós-pago</i>	51
Tabela 1B. <i>Modelo AR (7) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8</i>	53
Tabela 2B. <i>Modelo AR (7) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8</i>	53
Tabela 3B. <i>Modelo AR (7) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8</i>	54
Tabela 4B. <i>Modelo AR (7) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8</i>	54
Tabela 5B. <i>Modelo AR (4) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8</i>	54
Tabela 6B. <i>Modelo AR (2) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8</i>	55
Tabela 7B. <i>Modelo AR (4) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8</i>	55
Tabela 8B. <i>Modelo AR (7) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8</i>	55
Tabela 9B. <i>Modelo ARIMA (4, 1, 2) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8</i>	56

SUMÁRIO

<i>Capítulo I</i>	1
1. INTRODUÇÃO	1
1.1 <i>Breve historial da Empresa Telecomunicações de Moçambique, SARL:</i>	1
1.2 <i>Justificativa</i>	2
1.3 <i>Definição do problema</i>	2
1.4 <i>Objectivos</i>	3
1.4.2 <i>Específicos</i>	3
1.5 <i>Estrutura do Trabalho</i>	3
<i>Capítulo II</i>	4
2 REVISÃO DA LITERATURA	4
2.1 <i>Introdução</i>	4
2.2 <i>Série Temporal</i>	6
2.3 <i>Ruído Branco</i>	10
2.4 <i>Processo estocástico estacionário</i>	11
2.5 <i>Função de Autocorrelação</i>	12
2.6 <i>Passeio Aleatório</i>	14
2.7 <i>Testes de Raiz Unitária</i>	17
2.8 <i>Modelos de Previsão</i>	19
2.9 <i>Modelos auto-regressivos (AR)</i>	19
2.10 <i>Modelos de média móvel (MA – “moving average”)</i>	20
2.11 <i>Modelos mistos auto-regressivos e de média móvel (ARMA)</i>	21
2.12 <i>Modelos auto-regressivos integrados e de média móvel (ARIMA)</i>	21
2.13 <i>Método dos Mínimos Quadrados</i>	26
<i>Capítulo III</i>	28
3 METODOLOGIA	28
3.1 <i>Fonte de dados</i>	28
3.2 <i>Metodologia de Box-Jenkins</i>	28
3.3 <i>Pressupostos</i>	29
3.4 <i>Algoritmo para estimar um modelo ARIMA</i>	30

<i>Capítulo IV</i>	34
4 RESULTADOS	34
4.1 <i>Caracterização da série histórica do volume de clientes</i>	34
4.2 <i>Teste de Raiz Unitária</i>	37
4.3 <i>Estimação de modelos</i>	39
4.3.1 <i>Análise de Correlação</i>	39
4.5 <i>Seleção de modelos</i>	42
4.6 <i>Previsão</i>	44
<i>Capítulo V</i>	46
5 CONCLUSÕES E RECOMENDAÇÕES	46
5.1 <i>Recomendações</i>	46
5.2 <i>Limitações</i>	46
6 BIBLIOGRAFIA	47
<i>ANEXO A: Dados do volume de clientes dos serviços telefônicos pós-pago profissionais</i>	51
<i>ANEXO B: Estimação de Modelos</i>	53

1. INTRODUÇÃO

Na actividade de uma empresa, a previsão de volume de clientes é um factor essencial na condução do negócio, seja na perspectiva tradicional da empresa monopolista, ou na perspectiva actual da competição em ambiente de mercado regulado.

O processo de previsão pode exercer-se sobre objectos diversos e sob formas variadas, ou tendo em conta horizontes temporais e espaciais muito distintos. Para certos objectivos de planeamento, pode ser necessária uma previsão da evolução do volume de clientes a longo prazo.

A previsão do volume de clientes duma empresa é de extrema importância para auxiliar na determinação de recursos, tanto humanos como de equipamentos, para um atendimento eficaz aos clientes ajudando, deste modo, aos gestores na tomada de decisões e na aplicação de melhor estratégia no sentido de garantir o crescimento do negócio.

A previsão do volume de clientes é elaborada utilizando-se técnicas quantitativas e qualitativas ou ainda uma mistura de ambas. Os métodos quantitativos utilizam dados para prever a procura em períodos futuros baseando-se em modelos matemáticos, enquanto que, os métodos qualitativos baseiam-se em opiniões de especialistas. Para este trabalho, foi usada uma das técnicas quantitativas mais difundidas, o modelo ARIMA de Box-Jenkins, para a análise dos dados históricos.

O objectivo de estudo do presente trabalho, é a previsão do volume de clientes no serviço telefónico fixo pós-pago residenciais na província de Maputo, da Empresa Telecomunicações de Moçambique, SARL.

1.1 Breve historial da Empresa Telecomunicações de Moçambique, SARL:

A Empresa Telecomunicações de Moçambique foi criada em 1981 pelo decreto 05/81, de 10 de Junho, como E.E., num contexto de economia de planificação centralizada a partir da separação dos Correios,

Telégrafos e Telefones (CTT). Em 1992 foi transformada em empresa pública, E.P., através do decreto 23/92, de 10 de Setembro, num contexto de economia de mercado. Em 2002 foi transformada em empresa de direito privado, SARL (Sociedade Anónima de Responsabilidade Limitada), pelo decreto 47/2002, de 26 de Dezembro, designação essa que vigora até ao momento.

A Empresa Telecomunicações de Moçambique, SARL tem vindo a oferecer vários produtos e serviços, tais como: Serviços de Comunicação de voz (Telefone fixo pos-pago, Telefone fixo pre-pago, Cartão pre-pago bla-bla, Fixo sem Fio (CDMA), Postos públicos assistidos, Postos públicos a cartão, Televotação, Linha Verde (800), INMARSAT, Voice Mail, serviços suplementares), Serviços de Dados e Internet (Circuitos Alugados, RDIS, TDM Evolution, Internet Dedicada, Internet Sem Fio, TDM BandaLarga) e Serviços de Transmissão de Imagem (Videoconferencia). (TDM, SARL – Direcção de Marketing, 2008)

1.2 Justificativa

A introdução, no mercado de comunicações moçambicano, de serviços de telefonia móvel, tem sido apontada por muitos, como sendo a hipótese forte para justificar a tendência negativa dos índices de procura dos serviços telefónicos das Telecomunicações de Moçambique, a partir do ano de 2001, daí que, a adopção duma política de previsão do volume de clientes nos serviços telefónicos pós-pago, torna-se de extrema importância para o planeamento futuro do negócio da empresa.

A motivação da escolha do tema centra-se na avaliação da aplicabilidade do modelo estatístico ARIMA para o caso das Telecomunicações de Moçambique, SARL.

1.3 Definição do problema

Havendo um leque de ferramentas estatísticas que podem tornar as decisões mais rápidas e precisas do ponto de vista de previsão complementando o planeamento, surge então a seguinte questão: *Como prever o volume de clientes no serviço telefónico pós-pago profissionais das Telecomunicações de Moçambique, SARL, na província de Maputo?*

1.4 Objectivos

1.4.1 Geral

Mostrar a viabilidade teórica e prática do modelo ARIMA na previsão do volume de clientes nos serviços telefónicos pós-pago profissionais das TDM, SARL, na província de Maputo para o período de Maio a Dezembro de 2008.

1.4.2 Específicos

- Identificar um modelo ARIMA que melhor se ajusta a série de evolução de clientes;
- Prever o volume de clientes no período de Maio a Dezembro de 2008, na província de Maputo

1.5 Estrutura do Trabalho

Este trabalho está estruturado em cinco capítulos, que constam do seguinte:

Capítulo I: é constituído pela Introdução apresentando o tema abordado, justificativa para a escolha do mesmo e os objetivos do trabalho;

Capítulo II: apresenta a revisão de literatura dos assuntos pertinentes ao tema principal da dissertação, com destaque a série temporais, sua classificação, descrição de alguns métodos e aplicabilidade na previsão de valores futuros;

Capítulo III: descreve-se o material e os métodos aplicados;

Capítulo IV: é registada a discussão e interpretação dos resultados obtidos;

Capítulo V: Conclusões.

Finalmente, tem-se: Referências Bibliográficas e Anexos.

2 REVISÃO DA LITERATURA

Neste capítulo introduz-se o principal conceito utilizado ao aplicar a metodologia de Box-Jenkins.

2.1 Introdução

Para Tubino (2000) e Moreira (1996) devem ser assumidas algumas pressuposições em relação aos métodos de previsão. A primeira é que as causas que estiveram presentes no passado, configurando a procura, terão a mesma influência no futuro. A segunda é que quanto maior o horizonte previsto maior é a chance de erro.

Como a necessidade de previsão vem de diferentes áreas como vendas, produção, planeamento, é necessário que uma empresa desenvolva múltiplos enfoques para prever eventos incertos e criar um sistema para previsão. Isto, por sua vez, requer que a organização possua conhecimento e habilidades que abranjam pelo menos: (MAKRIDAKIS *et al*, 1983)

- Identificação e definição do problema a ser previsto.
- Aplicação de métodos de previsão.
- Procedimentos para selecção do método apropriado para uma específica situação.
- Suporte organizacional para a aplicação e utilização formal de métodos de previsão.

Quanto à realização de previsões de procura elas podem ser realizadas utilizando métodos quantitativos, qualitativos ou combinações de ambos (PELLEGRINI e FOGLIATTO, 2001).

Para Arnold (1999), os métodos qualitativos “são projeções subjectivas baseadas no discernimento, na intuição e em opiniões informadas”. A abordagem para o processo de previsão, de modo geral, não é explícita, mas baseada na experiência.

Segundo Tubino (2000), as técnicas quantitativas de previsão são aquelas que utilizam como procedimento a análise dos dados passados de maneira objectiva.

Para Levine *et al.* (2000), os métodos de previsão quantitativa utilizam dados históricos com o objetivo de estudar acontecimentos do passado para melhor entender a estrutura básica dos dados, e daí, prover os meios necessários para se prever ocorrências futuras.

Makridakis *et al.* (1998), observa que as técnicas quantitativas de previsão podem ser aplicadas sob três condições:

1. Quando há disponibilidade de informações referentes aos dados passados;
2. Quando as informações referentes aos dados passados estão na forma de dados numéricos;
3. Quando se pressupõem que os padrões que influenciaram os dados no passado continuam influenciando no futuro.

Conforme Corrêa (2000), os métodos quantitativos proporcionam uma boa aproximação da realidade a curto prazo desde que devidamente empregues. Entretanto, as incertezas sempre existirão e, geralmente, são de duas fontes distintas: a primeira delas é o próprio mercado, que pode ser bastante instável; a segunda corresponde ao próprio método de previsão.

O conceito de previsão, ou seja, de curto ou de longo prazo, está relacionado com a variável independente dos dados. Ou seja, para uma previsão de longo prazo (um ano, por exemplo) os dados utilizados não serão em horas ou em dias, mas em meses. Do mesmo modo que, utilizando dados horários ou diários, a intenção é termos previsões de curto prazo como, por exemplo, previsão para uma semana.

As técnicas de previsão quantitativa variam consideravelmente, tendo sido desenvolvidas com vários propósitos distintos. Cada técnica possui características próprias, grau de precisão e custo de utilização, os quais devem ser considerados na escolha de um método específico (Makridakis *et al.*, 1998).

Para Moreira (1996), os métodos quantitativos utilizam modelos matemáticos para a determinação de valores futuros e podem ser subdivididos em:

- Relacionamentos causais – a procura está correlacionada com variáveis internas ou externas à empresa;
- Séries temporais – avaliação do comportamento de valores tomados em instantes específicos de tempo.

De acordo com Davis *et al.* (2001), as análises das séries temporais, de modo geral, devem ser utilizadas para o curto prazo. As previsões causais, ao seu tempo, devem ser utilizadas em situações de longo prazo.

Segundo Tubino (2000), as técnicas baseadas em séries temporais procuram modelar matematicamente a procura futura relacionando os dados passados do próprio produto com o tempo, ou seja, partem do princípio de que a procura futura será uma projeção do passado.

Segundo Moreira (1996), os métodos que utilizam séries temporais procuram identificar uma tendência nas observações da procura ao longo do tempo. Via de regra, as observações levantadas tem um espaçamento constante (dias, semanas, meses, anos, etc.). Um dos pressupostos desses métodos é que os valores projectados podem ser estimados com base nos valores passados.

2.2 Série Temporal

Para Milone e Angelini (1995) chama-se série temporal a todo conjunto de valores passíveis de ordenação cronológica.

O estudo de séries temporais trabalha com o comportamento dos dados no passado e presente, para que estes forneçam a ideia de possíveis variações dos dados no futuro. A previsão é o principal objetivo do uso de séries temporais (Safadi e Nogueira, 2004).

Para Ribeiro e Paula (2000), uma série temporal é definida como um conjunto de observações de uma dada variável, geralmente distribuídas de maneira equidistante no tempo, e que possuem como característica central a presença de uma dependência serial entre elas. A série é denotada por Z_t , onde $t = \{1, 2, 3, 4, \dots, n\}$ são os intervalos de amostragem e com função densidade de probabilidade $p(Z_t)$ para cada t .

Segundo Ehlers (2003), uma série temporal é uma colecção de observações feitas sequencialmente ao longo do tempo. A característica mais importante deste tipo de dados é que observações vizinhas são dependentes e o maior interesse de pesquisadores é analisar e modelar esta dependência.

A série temporal também pode ser vista como a realização de um processo estocástico, que é definido como uma sequência de observações regidas por leis probabilísticas. Isto significa que uma série temporal pode ser considerada como uma amostra de um determinado processo estocástico (Ribeiro e Paula 2000).

Morettin e Tolo (1987) mostram que, quando se analisa uma série temporal os objetivos básicos são:

- i) Modelação do fenómeno sob consideração;
- ii) Obtenção de conclusões em termos estatísticos;
- iii) Avaliação da adequação do modelo em termos de previsão.

Deste modo, a aplicação de séries temporais utiliza a avaliação dos dados históricos para obter modelos de previsão. O objecto fundamental da análise de uma série temporal é determinar suas componentes básicas buscando identificar um padrão de comportamento da série que possibilite fazer previsões (Makridakis *et al.*, 1998).

Para a análise de qualquer série temporal, é essencial que, primeiramente, seja feito um gráfico, a fim de obter uma visão qualitativa geral do seu comportamento (FONSECA *et al.*, 1995). Entretanto, a inspeção visual do comportamento da série é apenas um direcionador sendo necessárias análises estatísticas para uma conclusão mais verdadeira sobre o comportamento da série.

Segundo Tubino (2000), as curvas obtidas pelas séries temporais podem conter tendências, sazonalidades, variações irregulares e variações randômicas.

- Tendência: indica a direção geral dos valores estudados. A principal característica deste componente é o movimento suave registrado num período longo de tempo, direcionando os dados de modo constante, crescente ou decrescente. Makridakis *et al.* (1998), asseguram que um padrão de tendência existe quando há um aumento ou diminuição do valor médio dos dados a longo prazo.
- As variações sazonais são flutuações que se repetem periodicamente, acompanhando um padrão temporal (relacionadas ao factor tempo), como por exemplo, influências climáticas ou a procura por algum produto ocorrem somente em determinado período do ano, onde se pode verificar a tendência de picos de vendas. A sazonalidade refere-se às mudanças ou variações cíclicas de curto prazo. São bastante parecidos com os cíclicos, com a diferença de que os fenômenos cíclicos são caracterizados por variações que oscilam em torno da tendência a intervalos aproximadamente regulares de tempo e em longo prazo. Porém, alguns autores preferem não fazer diferença entre os ciclos e a sazonalidade (ex: Moretin e Toloï, 1995). À luz de Levine *et al.* (2000), o ciclo “demonstra oscilações ou movimentos para cima e para baixo ao longo da série”.
- As variações irregulares, para Tubino (2000), são alterações na procura passada resultante de factores excepcionais, como greves ou catástrofes climáticas, que não podem ser previstos e, portanto, incluídos no modelo.
- Aleatoriedade (ou Erro): Também chamado de erro residual possui uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d)¹, frequentemente com média zero e variância constante igual a μ_2 São caracterizados por sua duração curta e intensidade variável. Para Mentzer e Bienstock (1998) esta componente da série temporal é denominada “ruído”.

¹ O pressuposto de aleatoriedade e independência refere-se à diferença entre cada valor observado e a média aritmética do seu próprio grupo, onde estas diferenças deveriam ser independentes para cada valor observado. Isso significa que a diferença (ou erro) para uma observação não deve corresponder à diferença (ou erro) de qualquer outra observação (Levine *et al.*, 2000).

Neste caso, os dados passados apresentam variações aleatórias ou randômicas. Estas variações, por sua vez, não podem ser modeladas pelas técnicas de previsão e são, de modo geral, tratadas pelas médias e desvio padrão ou variâncias.

Um outro tipo de classificação das séries temporais pode ser feito baseando-se no número de séries temporais envolvidas na modelagem. Esta classificação divide-se segundo Makridakis *et al.* (1998) em:

Modelos Univariados: grupo de dados que inclui todos os métodos que se baseiam em uma única série histórica, ou seja, a série temporal é explicada (prevista) apenas por seus valores passados. Este será o modelo utilizado no presente trabalho.

Modelos Multivariados ou Causais: grupo de dados que modela simultaneamente duas ou mais séries temporais permitindo, assim, relações de interdependência e causalidade, mas sem qualquer exigência com relação à direção da causalidade entre elas. A série temporal é explicada (prevista) pelos seus valores passados e também pelos valores passados de outras variáveis.

Para descrever séries temporais são utilizados modelos estatísticos que podem ser classificados em duas classes, segundo o número de parâmetros envolvidos. (MORETTIN e TOLOI, 2004).

- a) **Modelos paramétricos:** possuem um número de parâmetros finito. Os modelos mais comumente usados são os modelos AR, ARMA, ARIMA e modelos não lineares.
- b) **Modelos não-paramétricos:** possuem um número infinito de parâmetros. Podem ser exemplificados pela função de auto-covariância (ou autocorrelação) e modelos de redes neurais.

2.3 Ruído Branco

O termo erro é chamado de “ruído branco” se ele possui média zero, variância constante σ^2 , e é não autocorrelacionado. Ao definir-se um modelo de regressão, é conveniente assumir que o erro seja um ruído branco (Bryson e Ho, 1975).

A inclusão de uma perturbação (aleatória) ou termo-erro (com propriedades probabilísticas bem definidas) no modelo de regressão é necessária na análise de regressão por três importantes razões: (Stengel, 1986).

1. Desde que o propósito do modelo de regressão seja o de generalizar e simplificar. Processos económicos geralmente incluem somente os termos determinísticos de primeira ordem para o estudo. Isto significa que outras variáveis com efeitos de segunda ordem ou superiores não incluídas acabam por serem incluídas no termo erro, que pode ser visto como representante do efeito dos termos de segunda ordem e de ordens superiores;

2. A fim de levar em consideração o efeito de possíveis erros de medida da variável dependente ou da variável a ser explicada;

3. Desde que o factor humano difira de uma forma aleatória sob circunstâncias idênticas, o termo-erro pode ser usado para captar este factor humano inerentemente aleatório.

Para um bom ajuste dos modelos é necessário utilizar técnicas em que a estrutura residual seja um ruído branco, isto é, que o erro seja uma variável aleatória independente e identicamente distribuída, com distribuição normal, média zero e desvio padrão $\sigma - N(0;\sigma)$ (Sáfadi, 2003).

Caso contrário, se o erro tem um comportamento disperso, os modelos de previsão não conseguem modelar somente os dados e acabam por modelar também o erro e assim gerando instabilidades nas estimações.

Ehlers (2003), sugere um teste para verificar se o resíduo é independente e identicamente distribuído, isto é, se o resíduo é um ruído branco, é o teste de Box-Pierce-Priestley, o qual é baseado nas primeiras autocorrelações dos erros residuais.

2.4 Processo estocástico estacionário

Processo Estocástico (PE): é uma sequência de variáveis aleatórias x_t definidas em um espaço de probabilidades (Stengel, 1986).

Processo Estacionário: se as características estatísticas de um PE não se alteram no decorrer do tempo, os dados flutuam ao redor de uma média constante diz-se que o processo é estacionário (Makridakis *et al.*, 1998).

De acordo com Gujarati (2006), um processo estocástico é um conjunto de variáveis aleatórias ordenadas no tempo. E, define um processo estocástico estacionário como sendo aquele que 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, e não do próprio tempo em que a covariância é calculada.

É possível definir o carácter estacionário de uma série em termos matemáticos rigorosos, mas para o nosso propósito interessa-nos verificar que a série tem um aspecto plano, sem exibir tendência crescente ou decrescente, com dispersão regular em torno da média e sem flutuações periódicas indicando sazonalidade² (Menezes, 2007).

Assim um processo diz-se fracamente estacionário ou estacionário em covariância se:

$$\begin{aligned} E(X_t) &= \mu & \forall t = 1, \dots, T \\ \text{Cov}(X_t, X_{t-k}) &= \gamma_k & \forall t = 1, \dots, T \wedge k = 0, 1, 2, \dots \end{aligned} \quad (2)$$

² Denomina-se estacionariedade fraca

onde μ é uma constante real e γ_k tem uma estrutura constante para cada k .

2.5 Função de Autocorrelação

Por definição, autocorrelação é a correlação existente entre dois valores da mesma variável nos instantes x_i e x_{i+k} . Onde k é chamado de atraso (*lag*). (Box e Jenkins, 1976).

Autocorrelação: possui a finalidade de medir a intensidade com que um valor observado no tempo t é influenciado por aquele observado no tempo $t - k$.

Fazendo, na equação (2), $k = 0$ obtém-se a variância de X_t , dada por $Var(X_t) = \gamma_0$. O termo γ_k é uma função de k e é conhecido na literatura como a *função de auto-covariância*. Por outro lado, $\rho_k = \rho_{-k} = \gamma_k / \gamma_0$ é habitualmente designado por *função de auto-correlação* (do inglês *Autocorrelation Function* ou ACF) que desempenha, a par da média e da variância, um papel crucial no estudo da estacionariedade de um processo estocástico, uma vez que permite capturar a extensão temporal e robustez da memória do processo, ao medir a correlação dos valores actuais do processo com os seus valores passados. Note-se que

$$\rho_k = \frac{Cov(X_t, X_{t-k})}{Var(X_t)} = \frac{\gamma_k}{\gamma_0}, \quad (3)$$

onde $\gamma_0 = \sigma^2$ é uma constante, no pressuposto de que o processo X_t , $t = 1, 2, \dots, T$, é homoscedástico. Uma propriedade importante da função de auto-correlação (condição necessária) é que ela é uma forma semi-definida positiva, ou seja:

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \rho_{|i-j|} \geq 0, \quad (4)$$

para qualquer conjunto de momentos temporais t_1, t_2, \dots, t_n e quaisquer números reais $\alpha_1, \alpha_2, \dots$ (ver, e.g., Wei, 1990: 10-11).

A função de auto-correlação tem a propriedade de se atenuar à medida que k aumenta e apresenta comportamentos semelhantes em muitos casos sendo, por isso, difícil distinguir entre processos de ordem diferente. Para ajudar nessa distinção utiliza-se normalmente a chamada função de auto-correlação parcial (do inglês *Partial Autocorrelation Function* ou PACF) que se obtém a partir do seguinte processo AR (k):

$$X_t = \phi_{k1}X_{t-1} + \dots + \phi_{kk}X_{t-k} + a_t. \quad (5)$$

A k -ésima auto-correlação parcial é dada pelo coeficiente ϕ_{kk} do modelo (5). Uma característica importante da função de auto-correlação 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 (ver, e.g., Enders (2004: 65-67), para mais detalhes).

A função de autocorrelação (*Autocorrelation Function* – ACF) é usada para identificar a ordem de um modelo autoregressivo. Note que é necessário usar outra função, além de ACF, para identificar a ordem de um modelo autoregressivo, pois todos os modelos autoregressivos têm uma ACF que decresce exponencialmente e, portanto apenas o correlograma não traz informação sobre o grau do polinômio autoregressivo na representação autoregressiva com média móvel.

Em dados de séries temporais, o valor de uma variável em um período está normalmente correlacionado com seu valor no período seguinte. A correlação de uma série com seus próprios valores defasados é chamada de autocorrelação ou correlação serial. (Stock e Watson, 2004).

A forma da função de autocorrelação define o modelo a ser utilizado para o ajuste através da modelagem Box-Jenkins, além de determinar sazonalidade na série temporal e o período de ocorrência (Morettin e Toloí, 1985).

A função de autocorrelação parcial (*Partial Autocorrelation Function* – PACF) mede a correlação entre duas amostras x_t e x_{t+k} , separadas por um intervalo de tempo k , excluindo a dependência dos valores intermediários $w_{t+1}, w_{t+2}, \dots, w_{t+k-1}$.

Um caso particular de um processo fracamente estacionário é o processo ruído branco (do inglês *White Noise* ou WN). Num processo deste tipo a média é nula e a variância é finita: $\mu = \gamma_k = 0$ ($k \neq 0$) e $\gamma_0 = \sigma^2$. Contudo a média condicionada pela informação passada, $E(X_t | I_{t-1})$, onde I_{t-1} denota a história do processo até ao momento $t-1$, não é necessariamente nula em todos os passos do processo. Por este motivo é possível efectuar uma previsão residual num processo ruído branco, apesar dele não ser linearmente previsível (ver, e.g., Hendry, 1996).

2.6 Passeio Aleatório

Embora nosso interesse seja as séries temporais estacionárias, com frequência encontramos séries temporais não-estacionárias, dentre as quais o exemplo clássico é o modelo de passeio aleatório que é representado por

$$X_t = \mu + \rho X_{t-1} + \sigma \varepsilon_t, \quad (6)$$

onde $\varepsilon_t \sim \text{iid}(0, 1)$, isto é, é uma sequência de variáveis aleatórias independentes e identicamente distribuídas com média zero e variância unitária. Na equação (10) o termo μ é uma constante e o termo σ designa-se por *volatilidade*, ou seja, é o desvio-padrão dos resíduos. O coeficiente de autoregressão ρ determina a estabilidade do processo. Se $\rho = \sigma = 1$ e $\mu = 0$, o processo diz-se passeio aleatório. Para um valor inicial X_0 , o processo AR(1) é gerado pelos ruídos iid ε_t . Calculando as expressões para X_{t-1}, X_{t-2}, \dots , e substituindo recursivamente em (10), obtém-se:

$$X_t = \rho^t X_0 + \sum_{k=0}^{t-1} \rho^k (\mu + \sigma \varepsilon_{t-k}), \quad (7)$$

Donde se pode concluir que X_{t-1} é uma função linear de X_0 e de $\{\varepsilon_s; s = 1, \dots, t-1\}$, logo é independente de ε_t . Neste caso, os ε_t também se costumam designar por inovações ou processo de

inovação e corresponde à componente não previsível do modelo. Do mesmo modo, aplicando as condições dadas em (6) ao modelo AR(1) definido em (10), obtém-se:

$$\begin{aligned} E(X_t) &= \mu + \rho E(X_{t-1}) \\ \text{Var}(X_t) &= \rho^2 \text{Var}(X_{t-1}) + \sigma^2 \text{Var}(\varepsilon_t) \end{aligned} \quad (8)$$

Na equação (12) μ e σ^2 são constantes, $\text{Var}(\varepsilon_t) = 1$ e ρ é o parâmetro de estabilidade do processo, conforme anteriormente indicado. Assim, se $|\rho| \geq 1$, a variância de X_t é sempre maior do que a variância de X_{t-1} , logo X_t não pode ser um processo estacionário. Se $|\rho| > 1$, o processo diz-se explosivo e $X_t \rightarrow \infty$ com probabilidade 1. Se $|\rho| = 1$, sendo $\mu \neq 0$, o processo X_t designa-se por passeio aleatório com deriva. Neste caso, substituindo $\rho = 1$ em (7), obtém-se:

$$X_t = X_0 + \mu t + \sigma \sum_{k=1}^t \varepsilon_k, \quad (9)$$

Donde se conclui que o passeio aleatório com deriva tem uma tendência linear quer na média quer na variância, crescente na média se $\mu > 0$ e decrescente se $\mu < 0$ (ver, *e.g.*, Brockwell e Davis, 2002):

$$\begin{aligned} E(X_t) &= X_0 + \mu t \\ \text{Var}(X_t) &= t\sigma^2 \end{aligned} \quad (10)$$

O processo passeio aleatório também pode ser visto como uma soma de variáveis aleatórias iid. Basta notar que fazendo na equação (9) $X_0 = \mu = 0$ e $\sigma = 1$, obtém-se:

$$X_t = \sum_{k=1}^t \varepsilon_k = \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t. \quad (11)$$

Se $|\rho| < 1$, o processo X_t é estacionário e admite a seguinte solução estacionária para a média e variância:

$$\begin{aligned} E(X_t) &= \frac{\mu}{1-\rho} \\ \text{Var}(X_t) &= \frac{\sigma^2}{1-\rho^2} \end{aligned} \quad (12)$$

Num processo estacionário, as amplitudes dos choques ocorridos no sistema num dado momento vão-se atenuando ao longo do tempo, fazendo que o sistema regresse à situação de equilíbrio. Contudo, num processo não estacionário, as amplitudes dos choques são persistentes, não permitindo o retorno à situação de equilíbrio inicial. Neste caso, o carácter do processo vai sofrendo mutações irreversíveis ao longo do tempo.

O passeio aleatório com deriva pertence a uma classe de modelos designados por processos estacionários em diferenças (do inglês *Difference Stationary Processes* ou DSP) pois, diferenciando uma vez a variável original, obtém-se uma série que é estacionária, ou seja:

$$\Delta X_t = \mu + \sigma \varepsilon_t, \quad (13)$$

onde Δ representa o operador da primeira diferença, tendo os restantes símbolos sido definidos anteriormente. Deste modo, o passeio aleatório também se costuma designar por processo integrado de primeira ordem, sendo que $X_t \sim I(1)$ e $\Delta X_t \sim I(0)$. Diz-se, então, que X_t tem uma tendência estocástica.

Um outro processo não estacionário frequentemente encontrado em séries financeiras pertence à classe dos processos estacionários em tendência (do inglês *Trend Stationary Processes* ou TSP). Neste caso a componente não estacionária do processo pode ser capturada por um termo determinístico do tipo $\mu_0 + \mu_1 t$:

$$X_t = \mu_0 + \mu_1 t + \varepsilon_t, \quad (14)$$

Onde μ_0 e μ_1 são parâmetros. Este processo é estacionário em torno de uma tendência linear podendo, no entanto, ser usada qualquer outra função linearizável. Na prática, porém, é costume observar séries que contêm quer tendências determinísticas quer tendências estocásticas, pelo que uma possível estratégia de investigação passa por combinar os dois tipos de tendência acima referidos numa única equação.

A discussão anterior revela, embora de uma forma sucinta, a importância da determinação da ordem de integração de uma série observada x_t para efeitos de inferência sobre o processo gerador dos dados.

2.7 Testes de Raiz Unitária

Existem na literatura inúmeros testes que permitem analisar a questão da ordem de integração de uma série observada x_t . Estes testes são geralmente conhecidos como testes de não estacionariedade ou teste de raiz unitária, termo utilizado por analogia com a determinação das raízes da equação característica de uma equação com diferenças. Um dos testes de raiz unitária mais popular entre os investigadores é o teste Aumentado de Dickey-Fuller (do inglês *Augmented Dickey-Fuller test* ou ADF). O teste ADF, também conhecido por teste τ (Dickey e Fuller, 1979, 1981), baseia-se na seguinte regressão para a variável x_t :

$$\Delta x_t = \mu_0 + \mu_1 t + (\rho - 1)x_{t-1} + \sum_{k=1}^p \gamma_k \Delta x_{t-k} + \varepsilon_t, \quad (15)$$

Sendo a hipótese nula $\rho = 1$, usando para este efeito os valores críticos de MacKinnon (1991, 1996). O número de *lags* ou defasamentos do modelo é escolhido de modo a que os resíduos $\varepsilon_t \sim \text{iid}(0, \sigma^2_\varepsilon)$. Note-se que μ_0 é uma constante, $\mu_1 t$ captura uma eventual tendência determinística existente nos dados, conforme anteriormente referido e $(\rho - 1)x_{t-1}$ modela a respectiva tendência estocástica. Fazendo $\mu_1 = \gamma_k = 0$, a equação de Dickey-Fuller (15) reduz-se a um processo AR(1) que pode ser visto como uma realização do processo descrito na equação (6).

O procedimento usual nos testes ADF consiste em testar a hipótese nula $\rho = 1$ (existência de uma raiz unitária) para a variável em níveis (x_t) e em primeiras diferenças (Δx_t), contra a hipótese alternativa $|\rho| < 1$ (estacionariedade). Este teste pode considerar a existência de uma constante e uma tendência linear (τ_τ), de uma constante (τ_μ) ou de nenhuma delas (τ), enquanto componentes determinísticas do processo ($\mu_0 + \mu_1 t$). Para além do termo que contém o coeficiente de auto-regressão ρ , o modelo (15) incorpora ainda a possibilidade de existirem termos auto-regressivos da variável dependente Δx_t . Se a hipótese nula não for rejeitada em níveis, mas o for em primeiras diferenças, então a variável $x_t \sim I(1)$.

Se a hipótese nula for apenas rejeitada em segundas diferenças, a variável $x_t \sim I(2)$ e assim sucessivamente.

Dickey e Fuller (1979, 1981) propuseram dois testes para analisar as hipóteses acima descritas. O primeiro baseia-se na distribuição limite da estimativa OLS de ρ . A estatística do teste ADF utilizada neste contexto obedece à forma tradicional dos testes t :

$$t_{\rho} = \frac{\hat{\rho} - 1}{S_{\hat{\rho} - 1}}, \quad (16)$$

Mas não é assintoticamente normal nem simétrica. Deste modo, os valores críticos usuais da distribuição t -Student não são válidos neste âmbito. Fuller (1976), Dickey e Fuller (1981) e mais tarde MacKinnon (1991, 1996) fornecem os valores críticos adequados a esta situação. A verdadeira distribuição dos testes de Dickey-Fuller sob a hipótese nula é dada por (ver Dufrénot e Mignon, 2002):

$$t_{\rho} \underset{(17)}{\sim} \frac{W^2(1) - 1}{2\sqrt{\int_0^1 W^2(s) ds}}, \quad (17)$$

Onde $W(\cdot)$ representa um processo de Wiener no intervalo $[0, 1]$. O segundo, baseia-se na distribuição da estatística $T(\rho - 1)$, onde T denota a dimensão da amostra.

É interessante notar que $\tau_r < \tau_{\mu} < \tau$ e que, para a estatística τ , os valores críticos do teste ADF pouco diferem dos valores críticos da distribuição Normal estandardizada. Deste modo, o uso indevido destes últimos (ou, indiferentemente, dos valores críticos da distribuição t -Student) conduziria a uma sobre-rejeição da hipótese nula, aumentando o problema há medida que se introduzem mais componentes determinísticas no modelo. Marques (1998: 282-286), descreve um conjunto de estratégias aconselháveis para o estudo do teste de raiz unitária de Dickey-Fuller baseadas nos artigos citados.

Um resultado importante relativo aos testes de Dickey-Fuller é que a distribuição assintótica da estatística t_τ é independente do número de defasamentos das primeiras diferenças incluídos na regressão ADF. Estes, contudo, são necessários para eliminar qualquer auto-correlação de ordem superior remanescente na variável residual e, deste modo, validar o pressuposto de resíduos $\varepsilon_t \sim iid$ que está na base do uso do método dos mínimos quadrados. Para, além disso, a inclusão no modelo de uma componente de médias móveis (do inglês *Moving Average* ou MA) com um número adequado de termos desfasados das primeiras diferenças, não altera a validade assintótica do teste ADF (Said e Dickey, 1984).

2.8 Modelos de Previsão

A previsão de demanda utilizando métodos quantitativos pode ser feita através de vários modelos de regressão. O emprego de cada modelo depende basicamente do comportamento da série temporal que se deseja analisar (PELLEGRINI e FOGLIATTO, 2001)

2.9 Modelos auto-regressivos (AR)

Num processo auto-regressivo, o valor presente da série temporal X_t é expresso linearmente em termos dos valores passados da série e da perturbação aleatória e ocorrendo no instante t . A ordem deste processo depende do valor mais antigo sobre o qual a regressão tem lugar. Num processo *auto-regressivo de ordem p* , o modelo pode escrever-se como:

$$X_t = \delta + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t$$

em que os vários α_i são constantes reais e a série ε_t apresenta distribuição normal de valores independentes.

Uma variável dependente x_t segue um processo auto-regressivo de primeira ordem quando o seu valor num dado momento t depende apenas do seu próprio valor no período anterior ($t-1$) e de um termo

aleatório (ruído branco), ε_t , não correlacionado com x_t , com média zero e variância constante (Stengel, 1986).

Um modelo autoregressivo é simplesmente uma regressão linear do valor corrente da série sobre um ou mais dos valores anteriores da série. Por isso, estes modelos podem ser criados usando a técnica dos mínimos quadrados, e têm uma interpretação fácil.

2.10 Modelos de média móvel (MA – “moving average”)

Para Makridakis *et al.* (1998), o método consiste em calcular a média das últimas n observações mais recentes. O valor encontrado, ao seu tempo, é considerado a previsão para o próximo período.

Para Stengel (1986) a representação em média móvel ocorre quando a observação actual de uma variável possa ser explicada através de uma soma ponderada de ruídos anteriores e de um ruído actual.

Para Gujarati (2006), num processo dito de média móvel, o valor presente da série expressa-se em função dos valores presente e passados das perturbações aleatórias, que formam uma série de ruído branco. A ordem deste processo depende do valor mais antigo da série de ruído branco considerado; para *um processo de média móvel de ordem q*, a série exprime-se como

$$X_t = \delta + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_p \varepsilon_{t-p}$$

em que os vários α_i são constantes reais. Isto é, um modelo de média móvel constrói-se como uma regressão linear do valor presente da série sobre as perturbações aleatórias de um ou mais valores anteriores da série. Admite-se que estas perturbações são geradas por uma mesma distribuição, habitualmente normal, de média e desvio padrão constante.

Este modelo difere do anterior na medida em que cada perturbação aleatória se propaga para os valores futuros da série. Por causa disto, é mais complicado efectuar-se um ajustamento desta série e não se pode usar um modelo linear de mínimos quadrados. Além disso, os modelos MA são e mais difícil interpretação que os AR.

2.11 Modelos mistos auto-regressivos e de média móvel (ARMA)

Box & Jenkins popularizaram uma técnica que combina as características dos modelos autoregressivos AR e de média móvel MA. Embora estes modelos fossem já conhecidos e investigados, a contribuição de Box & Jenkins foi o desenvolvimento de um processo eficaz e sistemático de identificação e estimação de modelos que pudessem incorporar em simultâneo ambas características.

O processo de Box e Jenkins dito de tipo ARMA combina os dois efeitos anteriores. Um *processo de ordem p,q* representa-se por:

$$X_t = \delta + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q}$$

O modelo de Box e Jenkins assume que a série é estacionária. Nalguns casos, subtrai-se a série do valor médio, para se trabalhar com uma série de média nula.

2.12 Modelos auto-regressivos integrados e de média móvel (ARIMA)

Os processos definidos anteriormente são estacionários. Significa isto que a média da série temporal e as covariâncias entre as suas observações não variam com o tempo.

Se o processo não for estacionário, terá que se proceder à sua estacionarização. A estacionarização de uma série temporal recomendada por Box e Jenkins consegue obter-se com operações de diferenciação, mesmo que aplicada sucessivas vezes.

À série estacionarizada aplica-se então um modelo ARMA. Assim, ARIMA corresponde então a um processo auto-regressivo integrado com média móvel. A designação “integrado” explica-se pelo facto de se reconstruir a série original a partir da série diferenciada, por uma operação de integração ou soma recursiva.

São conhecidos também como Modelos Box-Jenkins univariados, em homenagem aos dois investigadores que mais contribuições fizeram a suas construções e aplicações: (BOX *et al.*, 1994).

Os modelos ARIMA compreendem:

- Modelo Auto-Regressivo de ordem p : AR (p): usado quando há autocorrelação entre as observações;
- Modelo de Média Móvel de ordem q : MA (q): usado quando há autocorrelação entre os resíduos;
- Modelo Auto-Regressivo de Média Móvel: ARMA (p,q); usado quando há autocorrelação entre as observações e autocorrelação entre os resíduos;
- Modelo Auto-Regressivo Integrado e de Média Móvel: ARIMA (p,d,q): usado em séries não estacionárias.

O modelo ARIMA é um caso geral dos modelos propostos por Box e Jenkins (1976), o qual é apropriado para descrever séries não estacionárias, ou seja, séries que não possuem média constante no período de análise, nas quais os parâmetros quase sempre são pequenos. Na prática, geralmente as séries encontradas apresentam tendência e (ou) sazonalidade (Sáfadi, 2003).

Os modelos ARIMA podem possuir termos auto-regressivos (a variável dependente e os termos defasados no tempo) e podem ser integrados quando seus dados originais não são estacionários. Será necessário diferenciar a série original dos dados d vezes até obter uma série estacionária (sua média e variância sejam constantes no tempo). Posteriormente, a série obtida pode ser modelada por um processo ARMA (p,q) (Stengel, 1986).

Para um bom ajuste do modelo ARIMA, é necessário utilizar técnicas em que a estrutura residual seja um ruído branco, isto é, que o resíduo seja uma variável aleatória independente e identicamente distribuída (Sáfadi, 2003).

O modelo ARIMA transforma uma série não-estacionária em série estacionária, ou seja, é necessário que a série não apresente tendência e sazonalidade. Para verificar a presença desses factores, é necessário aplicar alguns testes apropriados.

Em seguida é feita a regressão linear tendo como variáveis independentes a própria série defasada no tempo. Também, a sazonalidade é eliminada através da aplicação de médias móveis. Os coeficientes podem ser estimados pelos métodos de Máxima Verossimilhança.

Através de uma análise das autocorrelações absolutas e parciais, define-se o modelo mais adequado (Makridakis *et al.* 1998).

A metodologia para se prever uma série temporal através dos modelos ARIMA consiste de quatro etapas (Gujarati, 2006):

Etapa 1 - Identificação: ou seja, descobrir os valores apropriados de p , d e q dentre as várias versões dos modelos de Box-Jenkins, sejam eles sazonais ou não, que melhor descrevem o comportamento da série.

O primeiro parâmetro a ser identificado é o grau de diferenciação d necessário à estabilização dos dados. Isto é feito através do diagrama da função de autocorrelação (ACF), no qual são apresentados os valores das autocorrelações em relação aos atrasos (ou *lags*, são intervalos de tempo em que ocorrem máximas correlações) k . Se as autocorrelações decrescerem de forma exponencial, realizam-se diferenciações na série, até que o diagrama apresente um corte abrupto para um valor qualquer de autocorrelação, quando a série será considerada estacionária (Makridakis *et al.* 1998).

Em geral, o número de diferenciações d é no máximo igual a 2. Necessita-se de $d = 1$ quando a série é não estacionária em relação ao processo, isto é, quando a média durante um período é uma e depois desse período ocorre uma mudança na média. E necessita-se de $d = 2$ quando a série é não estacionária também quanto à inclinação, isto é, quando a série oscila em uma direção em um certo período e depois muda de direção (Makridakis *et al.* 1998).

A identificação do modelo a ser estimado ocorre pelo comportamento das ACFs e PACFs e seus respectivos correlogramas. A ordem autorregressiva p é determinada pela verificação da função de autocorrelação parcial (PACP) da série estudada. Se a série for unicamente autorregressiva ARIMA $(p,d,0)$, sua função de autocorrelação rk sofrerá uma queda exponencial gradativa, conforme pode ser

visto na *Figura 2.1* e sua função de autocorrelação parcial f_{kk} sofrerá uma queda repentina após o *lag* k , conforme pode ser visto na *Figura 2.1*. Se não, efectua-se uma análise dos estimadores para verificar até que ordem de defasagem do correlograma desta função é estatisticamente significativa. Essa será sua ordem autorregressiva (Ribeiro e Paula, 2000).

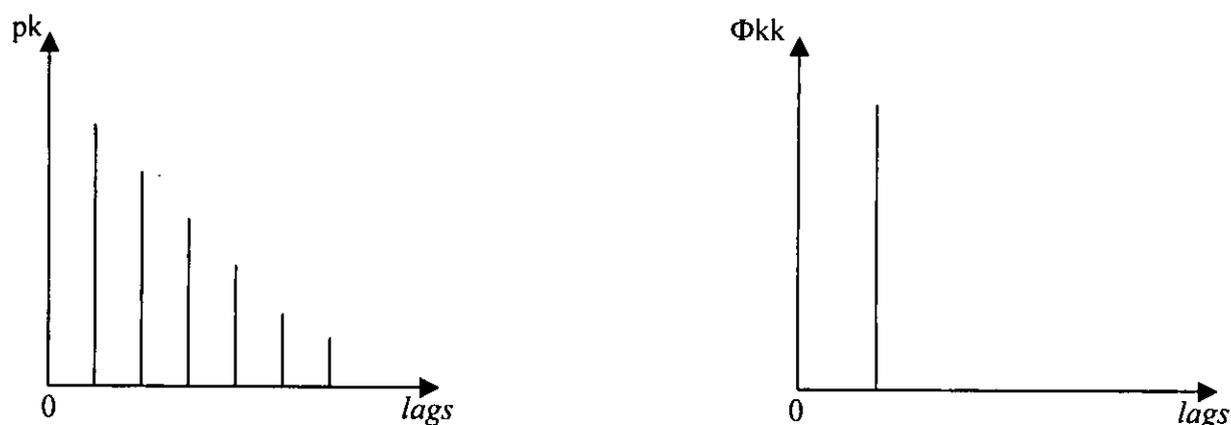


Figura 2.1 – ACF e PACF para um Modelo AR(1)

Se a série for unicamente de média móvel ARIMA (0,d,q), sua função de autocorrelação ρ_k sofrerá uma queda repentina após o *lag* k , conforme pode ser visto na *Figura 2.2* e sua função de autocorrelação parcial ϕ_{kk} sofrerá uma queda exponencial gradativa, conforme pode ser visto na *Figura 2.2*. (Ribeiro e Paula, 2000).

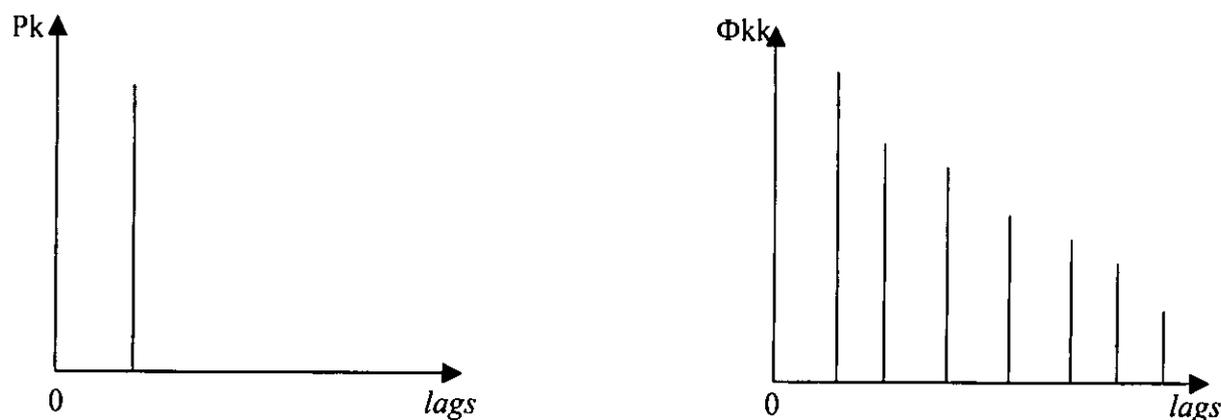


Figura 2.2 – ACF e PACF para um modelo MA (1)

Etapa 2 - Estimativa: depois de identificar os valores apropriados de p e q , o próximo passo é estimar os parâmetros dos termos auto-regressivo e de média móvel incluídos no modelo, que pode ser feito pelo método dos mínimos quadrados.

Etapa 3 - Verificação: depois de escolher um modelo ARIMA em particular, e estimar seus parâmetros, verifica-se se o modelo escolhido se ajusta aos dados razoavelmente bem, pois é possível que um outro modelo ARIMA possa desempenhar o mesmo papel. Um teste simples do modelo escolhido é ver se os resíduos estimados desse modelo são ruídos brancos; se forem, podemos aceitar o ajuste específico; se não são, devemos repetir o processo. Assim, a metodologia de Box e Jenkins é um processo iterativo.

Etapa 4 - Previsão: em muitos casos, as previsões obtidas com esse método são mais confiáveis do que as obtidas com a modelagem econométrica tradicional, especialmente para previsões a curto prazo. Naturalmente, é preciso checar cada caso.

Box e Jenkins (1976) formalizaram a teoria da utilização de componentes autorregressivos e de médias móveis na modelagem de séries temporais utilizando-se de duas idéias básicas na criação de sua metodologia de construção de modelos:

1. Parcimônia, que consiste na utilização do menor número possível de parâmetros para obter uma representação adequada no fenômeno em estudo.
2. Construção iterativa do modelo em que a informação empírica é analisada teoricamente sendo, o resultado deste estágio confrontado com a prática e assim sucessivamente até a obtenção de um modelo satisfatório.

O ciclo iterativo utilizado para a análise de uma série temporal através da metodologia de Box e Jenkins está representado na *Figura 2-3*

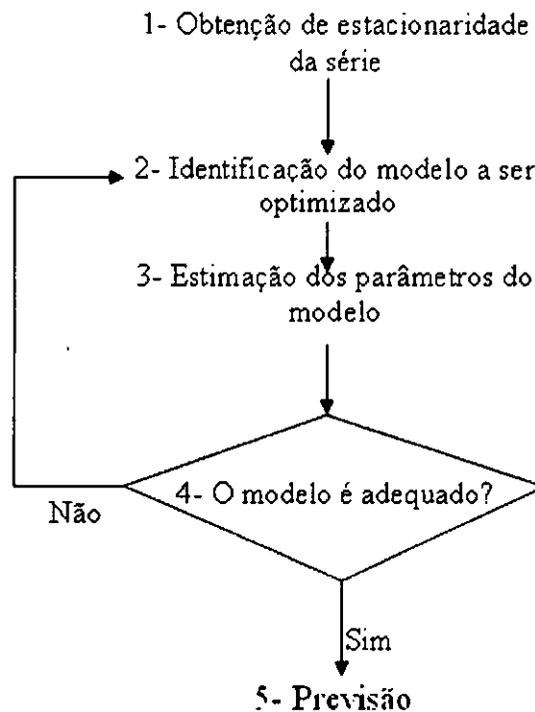


Figura 2-3 – Fluxograma do Modelo ARIMA de Box e Jenkins

Caso o modelo não seja adequado, o ciclo é repetido, voltando-se à fase de identificação *Figura 2-6*. Quando se obtém um modelo satisfatório, passa-se para a última etapa da metodologia de Box-Jenkins, que trata da realização de previsões.

2.13 Método dos Mínimos Quadrados

Gauss propôs estimar os parâmetros b_0 e b_1 na equação de modo a minimizar a soma dos quadrados dos desvios entre valores estimados e valores observados na amostra (ou seja, os desvios residuais). Este critério para estimar os coeficientes de regressão é chamado de método dos Mínimos Quadrados (Montgomery e Runger, 2003).

Para observações (X_i, Y_i) , temos o modelo:

$$Y_i = b_0 + b_1.X_i + e_i \quad i = 1, \dots, n$$

Onde: e_i é o ruído branco

b_0 é a interseção a ser estimada

b_1 é a inclinação do modelo a ser estimada

Deseja-se ajustar o modelo, estimando os parâmetros b_0 e b_1 .

O método dos Mínimos Quadrados considera os desvios de Y_i em relação ao seu valor esperado ($E(Y_i)$):

$$Y_i - (b_0 + b_1.X_i)$$

Elevando-se ao quadrado esses desvios e aplicando-se o somatório, temos o critério Q

$$Q = \sum (Y_i - b_0 - b_1.X_i)^2$$

De acordo com o método de Mínimos Quadrados, os coeficientes b_0 e b_1 são os valores estimados que minimizam o critério Q para a amostra $(X_1, Y_1), \dots, (X_n, Y_n)$ (Montgomery e Runger, 2003).

3 METODOLOGIA

Neste capítulo é apresentado todo material necessário para a realização do presente estudo. Faz-se mais uma vez, referência à metodologia de Box-Jenkins, como principal método de análise de dados, ilustrando-se as etapas para sua aplicação econométrica.

3.1 Fonte de dados

A variável utilizada no presente trabalho é a série histórica de volume de clientes do serviço telefónico pós-pago profissionais das TDM, SARL, na província de Maputo, no período de 2001 a Agosto de 2008 fornecida pela Direcção Comercial das TDM.

Para a materialização dos objectivos acima traçados, usa-se uma metodologia bastante conhecida no meio científico, o modelo ARIMA (*Auto-Regressivo Integrado de Média Móvel*) conhecido também como metodologia de Box - Jenkins.

3.2 Metodologia de Box-Jenkins

O objectivo da metodologia de Box-Jenkins (Box & Jenkins, 1970) é encontrar um modelo estocástico linear que possa ter gerado e que esse modelo possa ser utilizado para fornecer previsões de valores futuros da série. Um conceito importante nessa metodologia é o princípio da parcimónia. Tal princípio sugere que modelos mais simples, com poucos parâmetros, produzem melhores previsões que modelos super parametrizados. Um modelo parcimonioso ajusta bem os dados sem incorporar coeficientes inúteis e quanto mais baixos os valores das estatísticas de Akaike e de Schwarz, melhor o modelo. O objectivo é se aproximar do processo gerador original dos dados e não descrevê-lo exactamente.

Estacionaridade - O primeiro passo para desenvolver um modelo de Box - Jenkins é determinar se a série é estacionária. Isto se consegue observando se o gráfico da série apresenta uma escala constante e posicionamento estável. Também se pode observar, num diagrama de autocorrelação, quando a série não é estacionária, verifica-se um decaimento muito lento da função de autocorrelação.

Obtenção de estacionaridade: Box - Jenkins recomendaram o processo de diferenciação para se obter a estacionaridade de uma série, mas também é possível ajustar-se uma curva e trabalhar com a série das diferenças para esta curva.

Se a série não for estacionária, é habitualmente possível convertê-la numa outra estacionária recorrendo a algumas técnicas apropriadas:

- a) Se os dados revelam uma tendência ("trend") crescente ou decrescente, pode tentar-se ajustar uma curva e subtraí-la aos dados, ficando com uma série de resíduos. Como o propósito do ajustamento é remover a tendência de longo prazo, habitualmente um ajustamento de uma recta é suficiente;
- b) Se a variância não é constante, extrair o logaritmo ou a raiz quadrada da série pode ajudar a estabilizá-la.
- c) Em qualquer caso, pode-se diferenciar a série. Isto é, dada uma série com elementos X_j , constrói-se a série das diferenças Y_j

$$Y_j = X_j - X_{j-1}$$

A série diferenciada conterá menos um ponto que a original. Embora se possa diferenciar uma série mais do que uma vez, é comum que uma diferenciação seja suficiente. Como é evidente, a partir de uma série diferenciada pode reconstituir-se uma série original por integração, ou seja, pela soma acumulada dos elementos da série diferenciada.

3.3 Pressupostos

A fim de se obter melhores resultados na utilização da metodologia de Box-Jenkins (ARIMA), três hipóteses básicas precisam ser observadas. A primeira é relativa ao tamanho inicial da amostra geralmente aceita, que deve ser de, no mínimo, 50 observações (Box & Jenkins, 1976). Este tende a ser um obstáculo significativo para muitos dados que somente são colectados anualmente. A segunda

suposição é a de que a série de dados seja estacionária, isto é, que a série varie em torno de uma média constante e com uma variância constante.

Rodar uma regressão com duas variáveis não-estacionárias pode causar resultados espúrios (Granger & Newbold, 1974). Se os dados forem não-estacionários, a série de dados deve ser diferenciada ou deve-se proceder à adição de uma tendência no tempo. Se os dados forem de tendência não-estacionária somente, então, adicionar ao modelo uma tendência linear do tempo tornará a série estacionária. Os dados com tendência não-estacionária têm uma média e uma variância que mudam com o tempo, mas por uma quantidade constante. Se os dados forem não-estacionários em primeira diferença, uma única diferenciação tornará a série estacionária.

A terceira hipótese para os modelos ARIMA é a de que a série seja homoscedástica, isto é, tenha uma variância constante ao longo do tempo. Se a amplitude da variação em torno da média aumentar com o tempo, mesmo após a diferenciação, a série é considerada heteroscedástica. A solução para este problema pode ser simples ou complexa e envolve medidas de ajustes tais como: usar o logaritmo natural dos dados, usar raízes quadradas ou cúbicas, truncar a série de dados, entre outros.

O primeiro componente do processo ARIMA é o auto-regressivo. Tal componente prevê os valores futuros com base em uma combinação linear dos valores passados. O segundo componente, de médias móveis, fornece as estimativas das previsões com base nos erros presentes e passados de previsão. Estes dois componentes formam, juntos, os modelos auto-regressivo de médias móveis (ARMA). Se uma série possuir tendência ou diferença estacionária, o modelo se transforma no modelo ARIMA de Box-Jenkins (Box & Jenkins, 1976).

3.4 Algoritmo para estimar um modelo ARIMA

A modelagem ARIMA se compõe, basicamente, de quatro estágios, a saber: Identificação; Estimação; Diagnóstico e Previsão.

Identificação

Segundo Box, G.E.P., Jenkins (1976), Esta etapa consiste em descobrir os valores apropriados de p (ordem da componente auto-regressivo), d (grau de diferenciação da série) e q (ordem da componente de média móvel), através da análise da função auto-correlação (FAC) e auto-correlação parcial no correlograma.

As principais ferramentas da identificação são: a função auto correlação (FAC), a função autocorrelação parcial (FACP) e os correlogramas resultantes, que são simplesmente as representações gráficas das FACs e FACP contra o tamanho da defasagem. Para o uso desta etapa é obrigatório conhecer os modelos AR, MA e ARMA por meio das características da FAC e da FACP.

Identificação das ordens p e q - isto se executa pelo exame das funções de autocorrelação e de autocorrelação parcial, comparando os seus gráficos com o respectivo comportamento teórico quando as ordens p e q são conhecidas.

Ordem p de um processo auto-regressivo - para um processo auto-regressivo AR(1) de ordem 1, a função de auto-correlação deve ter um comportamento com decaimento exponencial. Para processos AR de ordem superior, poderão observar-se combinações de aspectos de decaimento exponencial com sinusóides amortecidas. Neste caso, interessa observar a função de autocorrelação parcial, cujo valor para um processo AR (p) deverá ser nulo para atrasos superiores a p .

Ordem q de um processo de média móvel - um processo de média móvel MA pode ser indiciado por uma função de autocorrelação com um ou mais picos. A função de autocorrelação de um processo MA (q) anula-se para atrasos superiores a q .

A tabela seguinte resume como usar a função de autocorrelação para identificar o modelo de uma série temporal:

Tabela 3-1: . Padrões teóricos das funções de autocorrelação e auto-correlação parcial

Forma	Modelo indicado
Exponencial, decaindo para zero	Modelo AR. Usar o gráfico da função de autocorrelação parcial para identificar a ordem p do processo
Valores alternando entre positivos e negativos, decaindo para zero	Modelo AR. Usar o gráfico da função de autocorrelação parcial para identificar a ordem p do processo
Um ou mais picos, com os restantes valores praticamente zero	Modelo MA. A ordem q do modelo é identificada pelo valor do atraso em que a função se anula
Decaimento só começando após alguns atrasos	Modelo misto ARMA

Fonte: Ehlers (2003)

Estimação de parâmetros

Uma vez identificado o modelo (AR, MA, ARMA, ARIMA), procede-se à estimação dos seus parâmetros, recorrendo-se em geral a algum processo ou algoritmo que minimiza os quadrados dos resíduos ou no princípio da máxima verosimilhança. Não desenvolveremos nestas notas introdutórias tais processos.

Diagnóstico do modelo

Depois de escolher um modelo ARMA em particular, e estimar seus parâmetros, vemos em seguida se o modelo escolhido se ajusta aos dados razoavelmente bem, pois é possível que um outro modelo ARMA possa ajustar melhor. Um teste simples do modelo escolhido é ver se os resíduos estimados desse modelo são de ruído branco; se forem, pode-se aceitar o ajuste específico; se não são ruído branco, deve-se retornar ao passo 1. Assim, a metodologia BJ é um processo iterativo.

Previsão

Em muitos casos, as previsões obtidas com esse método são mais confiáveis do que as obtidas com a modelagem econométrica tradicional, especialmente para previsões de curto prazo (Gujarati, 2006)

Em modelos ARMA as previsões podem ser obtidas usando-se directamente a equação do modelo. Assumindo que a equação do modelo seja conhecida a previsão $\hat{x}_n(k)$ é obtida substituindo valores futuros dos erros ε por zero, valores futuros da série X_{n+1}, X_{n+2}, \dots , pela sua esperança condicional, e valores passados de X e de ε pelos seus valores observados.

Os intervalos de confiança 95% são construídos para as estimativas dos pontos preditos pelo modelo, que são calculados assumindo a normalidade dos erros-padrão das estimativas, sendo que neste caso, para uma série temporal X_i :

Resumindo, Pode-se notar que nesta metodologia os dados tem de ser submetidos a um ciclo iterativo, cujo o processo consiste em submete-los a 4 (quatro) etapas distintas, que são: identificação, estimação, verificação e previsão, tal como se ilustra no diagrama da figura 2-3.

4 RESULTADOS

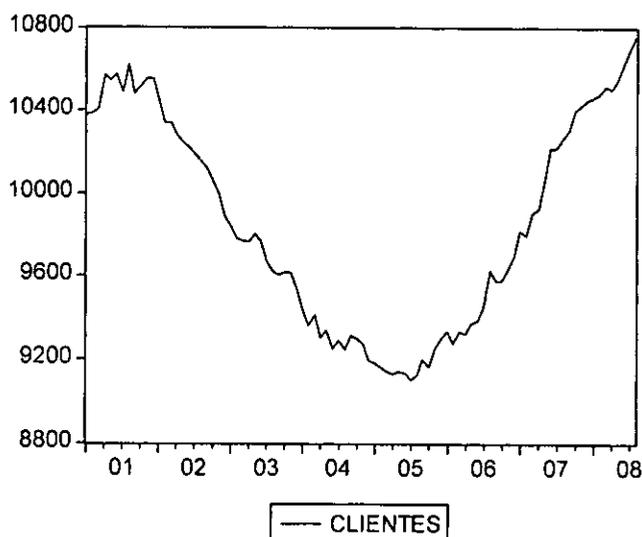
Neste capítulo apresenta-se a análise empírica do estudo. Inicialmente é feita a descrição das séries utilizadas. Em seguida são feitos testes de raiz unitária para verificar a estacionaridade da série e por fim usa-se a metodologia de Box-Jenkins descrita no capítulo 3 para encontrar um modelo consistente e representativo para a previsão do volume de clientes das TDM, SARL.

4.1 Caracterização da série histórica do volume de clientes

Neste estudo, utilizou-se uma série de dados mensais referentes ao volume de clientes que aderiram aos serviços telefónicos pós-pago profissionais das TDM, SARL, na província de Maputo, de Janeiro de 2001 a Agosto de 2008, totalizando 92 observações.

Uma representação gráfica dos dados é, em geral, o primeiro passo na análise de qualquer série temporal. A Figura 4-1 mostra o gráfico da evolução temporal do volume de clientes dos serviços telefónicos pós-pago profissionais das TDM, SARL utilizada neste estudo. Pela observação da evolução da série, constata-se a existência de dois sub-períodos, um até 2005 e outro após 2005. De 2001 até 2005 nota-se uma considerável diminuição do número de clientes das TDM, mas de 2005 até 2008, nota-se que a procura por estes serviços volta a aumentar.

Figura 4-1. Evolução temporal do volume de clientes nos serviços telefónicos pós - pago profissionais das TDM, SARL em nível, 2001-1 a 2008-8.



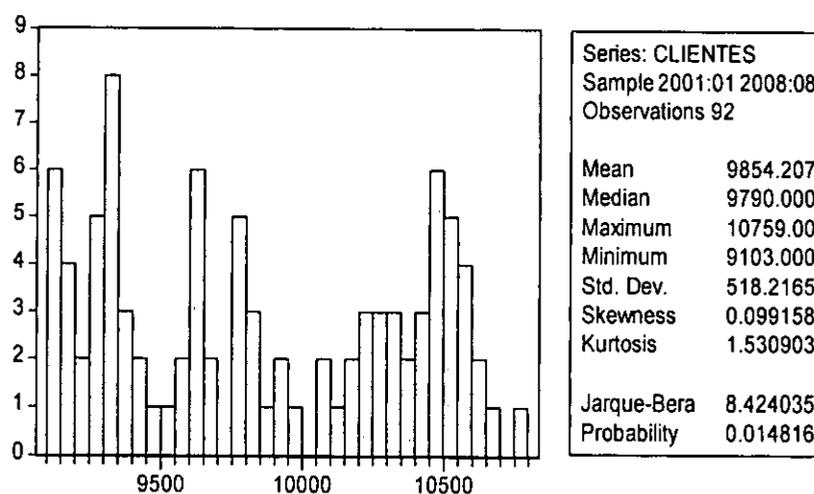
Fonte: Adaptação da autora

Admite-se que esta alteração estrutural na evolução temporal do volume de clientes nos serviços telefónicos pós-pago profissionais das TDM, SARL se deva:

- a) A entrada em vigor dos serviços da telefonia móveis em Moçambique, no ano de 1997, muitas empresas ficaram atraídos em experimentar aqueles novos serviços, em detrimento dos serviços de telefonia fixa.
- b) E a introdução em 2005, dos serviços de Internet BandaLarga na TDM, SARL acoplados às redes telefónicas fixas, levou a um rápido aumento na procura destes serviços.

No entanto, em ambos sub-períodos (antes e depois de 2005), a série em estudo apresentou uma tendência decrescente (até 2005) e crescente (depois de 2005) o que sugere que a série em estudo possui uma raiz unitária. Isto é, possui uma média e variância mudando ao longo do tempo.

Figura 4-2: Histograma do volume de clientes nos serviços telefónicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8.



Fonte: Adaptação da autora

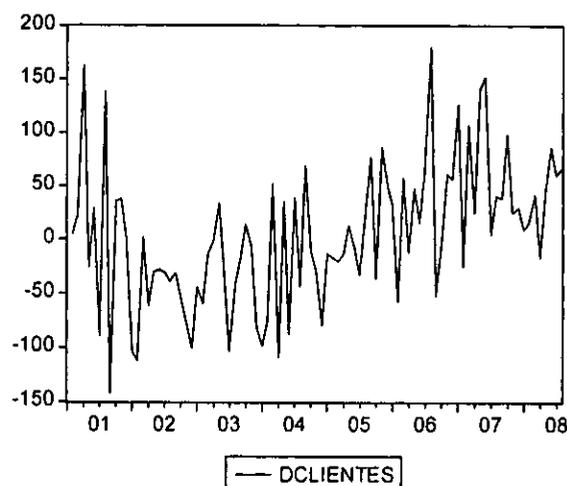
Pela análise da figura anterior e das respectivas estatísticas, constata-se que a série temporal em estudo apresenta uma assimetria com desvio à direita, com um *skewness* (coeficiente de assimetria) de cerca de 0.099158.³ O índice de *kurtosis* (coeficiente de achatamento), cerca de 1.530903 da série indica que se está perante uma série que não segue a distribuição normal⁴. O que é confirmado com a estatística de Jarque-Bera (8.42) associada a uma probabilidade de 0.014816 < 0.05, o que leva a rejeitar a hipótese de que as observações tem uma distribuição normal.

No entanto, a série do volume de clientes em primeiras diferenças não apresenta nenhuma tendência, assemelhando-se ao comportamento de uma série de ruído branco, embora haja períodos que apresenta grandes oscilações seguidas de períodos em que há relativa estabilidade, sugerindo que a série não tem variância constante ao longo do tempo (aglomeração da volatilidade).

³ Para uma curva normal simétrica o coeficiente é nulo

⁴ Para uma distribuição normal o kurtosis seria de 3

Figura 4-3. *Evolução temporal do volume de clientes nos serviços telefónicos pós-pago profissionais da TDM, SARL em primeiras diferenças, 2001-1 a 2008-8.*



Fonte: Adaptação da autora

4.2 Teste de Raiz Unitária

O processo de modelação das séries temporais financeiras implica que, em primeiro lugar, se proceda à análise das propriedades de estacionariedade da série em estudo. Para o efeito, utilizar-se-á o teste de raiz unitária de Dickey e Fuller descrito no capítulo 2. Nota-se que este teste permite detectar a existência ou não de estacionariedade na série em estudo. Os resultados encontram-se nas Tabelas 4-1 e 4-2.

As hipóteses to teste ADF estão descritas da seguinte forma:

H_0 : A série em estudo tem raiz unitária (não é estacionária)

H_1 : A série em estudo não tem raiz unitária (é estacionária)

Tabela 4-1. Teste de raiz unitária em nível com intercepto e tendência para o logaritmo do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8.

Null Hypothesis: LCLIENTES has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Automatic based on SIC, MAXLAG=11)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	0.943956	0.9998
Test critical values: 1% level	-4.062040	
5% level	-3.459950	
10% level	-3.156109	

*MacKinnon (1996) one-sided p-values.

Fonte: Adaptação da autora

A tabela acima (4-1), apresenta o teste de raiz unitária de Dickey-Fuller, o qual não se rejeita a hipótese nula de existência de uma raiz unitária na série em estudo, uma vez que a estatística de ADF (0.943956) é menor, em módulo, que o valor crítico de Mackinnon a 5%. Chega-se à mesma conclusão observando-se a probabilidade de 0.99 superior ao nível de significância de 5%. Contudo, este resultado era de esperar uma vez que a maioria das séries económicas é não estacionária em nível.

Portanto, A correcção da não-estacionaridade foi feita através da aplicação da transformação da função logarítmica para estabilização da variância e tomadas as diferenças para estabilização da média, e submeteu-se novamente ao teste de raiz unitária de Dickey-Fuller, conforme a Tabela 4-2.

Tabela 4-2. Teste de raiz unitária em primeiras diferenças com intercepto e tendência para volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8.

Null Hypothesis: D(LCLIENTES) has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Automatic based on SIC, MAXLAG=11)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-9.819439	0.0000
Test critical values: 1% level	-4.063233	
5% level	-3.460516	
10% level	-3.156439	

*MacKinnon (1996) one-sided p-values.

Fonte: Adaptação da autora

Desta vez, a estatística de ADF (-9.819439) é maior em módulo que o valor crítico ao nível de significância de 5%, portanto, rejeita-se a hipótese nula, o que significa que existem evidências

estatísticas suficientes para se afirmar que a série das primeiras diferenças do volume de clientes em estudo é estacionária.

4.3 Estimação de modelos

Para identificar e estimar o modelo mais adequado é necessário em primeiro lugar, fazer uma análise das funções de autocorrelações absolutas e parciais.

4.3.1 Análise de Correlação

Na *Figura 1A em anexo*, apresenta-se os correlogramas de autocorrelação e autocorrelação parcial da série do volume de clientes. Dois factos se destacam: primeiro, a função de autocorrelação decai muito lentamente, até a defasagem 9, as funções de auto-correlação são individualmente diferentes de zero, já que todas estão fora do intervalo de 95% de confiança. Segundo, após a primeira defasagem, a FAC cai substancialmente e todas as FAC que se seguem não são significativos do ponto de vista estatístico.

Nota-se que a FAC descreve um comportamento padrão de uma série de passeio aleatório, isto é, esta série é típica de um processo estocástico não estacionário o que foi confirmado com o teste de raiz unitária de ADF (vide figura 1A em anexo).

No entanto, o correlograma das funções de autocorrelação e da autocorrelação parcial da série em primeiras diferenças apresenta um padrão bem diferente. As FACP nas defasagens 2, 4, 5 e 7 parecem serem diferentes de zero em termos estatísticos, mas todas as demais defasagens não são estatisticamente diferentes de zero (vide figura 2A em anexo).

As autocorrelações parciais com picos nas defasagens 2, 4, 5 e 7 parecem estatisticamente significativas, mas as demais não são. Portanto, vai-se supor que o processo que gerou (a primeira diferença) do volume de clientes seja no máximo um processo AR (7). Obviamente, não tem que se incluir todos termos AR até 7, pois a partir do correlograma parcial sabe-se que apenas os termos AR das defasagens 2, 4, 5 e 7 são significativos.

Designa-se as primeiras diferenças da série do volume de clientes nos serviços telefónicos pós pago residenciais da TDM, SARL, 2001-1 a 2008-8 por Y_t^* , o modelo provisoriamente identificado como AR é:

$$Y_t^* = \delta + \alpha_1 Y_{t-2} + \alpha_2 Y_{t-4} + \alpha_3 Y_{t-5} + \alpha_4 Y_{t-7} \quad (4.1)$$

Com o auxílio do Eviews, obteve-se as estimativas da tabela 1B em anexo cujo modelo é representado por:

$$Y_t^* = 0.000664 + 0.201775Y_{t-2} + 0.265524Y_{t-4} + 0.129203Y_{t-5} + 0.099288Y_{t-7} \quad (4.2)$$

T (tau) = (0.195821) (1.982605) (2.658298) (1.339836) (2.241593)

$R^2=0.304713$ DW=1.8019

Analisando-se o modelo acima, constata-se que a constante e o termo AR (5) apresentam coeficientes não significativos em termos estatísticos. Portanto, é possível que exista um outro modelo que melhor se ajusta aos dados sem incluir os termos não significativos. Por isso, vai-se estimar um modelo sem intercepto e sem o termo AR (5) como a seguir se ilustra.

$$Y_t^* = \alpha_1 Y_{t-2} + \alpha_2 Y_{t-4} + \alpha_4 Y_{t-7} \quad (4.3)$$

Donde obteve-se as seguintes estimativas da tabela 2B em anexo

$$Y_t^* = 0.200167Y_{t-2} + 0.274989Y_{t-4} + 0.269661Y_{t-7} \quad (4.4)$$

T (tau) = (1.969037) (2.763575) (2.904959)

Verifica-se que, a um nível de significância de 5% o coeficiente do parâmetro AR(2) não é significativo (prob=0.0524>0.05).

Visto que o modelo 4.4 apresenta um dos seus coeficientes não significativos a 5%, vai-se achar um outro modelo ARIMA que tenha coeficientes significativos. Para tal vai-se eliminar o termo AR (2) estimando-se um modelo apenas com os termos 4 e 7 representado da seguinte forma:

$$Y_t^* = \alpha_1 Y_{t-4}^* + \alpha_2 Y_{t-7}^* \quad (4.5)$$

O modelo encontrado consta na tabela 3B em anexo, com a seguinte representação:

$$Y_t^* = 0.345857Y_{t-4}^* + 0.314365Y_{t-7}^* \quad (4.6)$$

$$T(\text{tau}) = (3.664463) \quad (3.432905)$$

Constata-se que os coeficientes são estatisticamente significativos, mas o R^2 vai diminuindo com a retirada de alguns termos auto-regressivos. Mas na verdade, isso não constitui problema porque sempre que se aumenta ou se diminui variáveis num modelo, o R^2 vai aumentar ou diminuir respectivamente, mesmo que estas variáveis não sejam significativas (Anderson, 2003).

Mesmo tendo-se encontrado um modelo que apresenta coeficientes estatisticamente significativos e que tenha um menor número de parâmetros, houve necessidade de se encontrar outros modelos com vista a melhorar o coeficiente de determinação R^2 , mas sem aumentar o número de parâmetros. Portanto, estimou-se outros modelos AR (7) apresentados nas tabelas 4B, 5B, 6B, 7B, 8B e 9B em anexo.

Quando se estimou o modelo AR (7) o coeficiente do termo AR (2) estava associado a um valor de p-value maior que 5% o que se levou a retirá-lo do modelo uma vez que não havia evidências estatísticas que suportam a sua inclusão. No entanto, estimou-se o modelo com os termos AR(2) e (7) e um outro com os termos AR(2) e (4). em ambos modelos, o coeficiente do termo AR(2) mostrou-se estatisticamente significativo. Os resultados da estimação constam nas tabelas 4B e 5B em anexo. Nota-se que o modelo 5B apresenta um R^2 maior que o da tabela 4B, mas ambos são superados com o do modelo com os termos AR (4) e (7) da tabela 3B em anexo.

Conforme foi descrito no capítulo 3, o algoritmo para encontrar um modelo de Box-jenkins é interativo. Por isso, houve necessidade de se encontrar outros modelos ARIMA que melhor se ajustam aos dados da série em estudo. Pelo que, estimou-se um modelo que tenha apenas o termo AR (2) outro com AR (4) e outro com AR (7). Os resultados constam nas tabelas 6B, 7B e 8B em anexo, respectivamente. Os três modelos apresentam coeficientes significativos embora com um coeficiente de determinação baixo.

Por último, estimou-se um modelo misto ARIMA (4,1,2) na tentativa de melhorar o grau de predição dos modelos. Os resultados constam na tabela 9B em anexo.

Os coeficientes do modelo são estatisticamente significativos e conseguiu-se aumentar o R^2 .

4.4 Diagnóstico dos modelos estimados

Primeiramente, o objectivo é verificar se o modelo é parcimonioso, isto é, se não tem parâmetros em excesso. A verificação da existência de parâmetros em excesso é feita com base no erro padrão dos coeficientes analisando o valor de p-value do parâmetro estimado, caso o valor de p-value seja maior que um nível de confiança pré-estabelecido (neste caso 5%), conclui-se que ele não é significativo, ou seja, não há evidências estatísticas para suportar a inclusão desse coeficiente no modelo.

Depois, gerou-se o resíduo de cada modelo e obtiveram-se as funções de autocorrelação e autocorrelação parcial desses resíduos até a defasagem 20. as FAC e FACP são apresentados no anexo C. Nenhuma das autocorrelações geral e parcial são estatisticamente significativas. Pelo que, dá se a impressão de que os resíduos de todos modelos estimados são de ruído branco. Portanto, todos os modelos se ajustam razoavelmente aos dados. Para saber qual desses modelos melhor se ajusta aos dados da série em estudo, vai-se utilizar os critérios de selecção de modelos descritos no capítulo 3.

4.5 Selecção de modelos

Para a selecção do melhor modelo, utiliza-se os critérios de AIC (Akaike Information Criteria) e o SBC (Schwartz Bayesian Criteria) descritos no capítulo 3.

Tabela 4-3: Estatísticas de estimação de modelos

Modelo	b_i	se	t	P-value	R^2	DW	AIC	SBC	Lglikeliho od	
1	C	0.000664	0.003389	0.195821	0.8453	0.305	1.80	-7.488	-7.344	319.5
	X(-2)	0.201775	0.101773	1.982605	0.0509					
	X(-4)	0.265524	0.099885	2.658298	0.0095					
	X(-5)	0.129203	0.096432	1.339836	0.1841					
	X(-7)	0.222564	0.099288	2.241593	0.0278					
2	X(-2)	0.200167	0.101657	1.969037	0.0524	0.288	1.82	-7.513	-7.426	318.5
	X(-4)	0.274989	0.099505	2.763575	0.0071					
	X(-7)	0.269661	0.092828	2.904959	0.0047					
3	X(-4)	0.345857	0.094381	3.664463	0.0004	0.254	1.82	-7.490	-7.432	316.6
	X(-7)	0.314365	0.091574	3.432905	0.0009					
4	X(-2)	0.301782	0.098536	3.062672	0.0030	0.221	1.74	-7.447	-7.389	314.8
	X(-7)	0.270404	0.096511	2.801784	0.0063					
5	X(-2)	0.266303	0.099668	2.671894	0.0090	0.230	1.99	-7.426	-7.369	325.0
	X(-4)	0.305656	0.100661	3.036476	0.0032					
6	X(-2)	0.375651	0.099726	3.766839	0.0003	0.136	1.80	-7.295	-7.267	325.0
7	X(-4)	0.402809	0.097157	4.145946	0.0001	0.166	1.96	-7.368	-7.340	321
8	X(-7)	0.348279	0.097688	3.565198	0.0006	0.132	1.71	-7.362	-7.333	310.2
9	X(-4)	0.373804	0.100990	3.701413	0.0004	0.224	1.92	-7.418	-7.361	324.6
	e(-2)	0.295017	0.106930	2.758979	0.0071					

Fonte: Adaptação da autora

Analisando-se a tabela acima, constata-se que dos 9 modelos estimados, todos descrevem numa forma parcimoniosa o comportamento do volume de clientes dos serviços telefónicos pós-pago residenciais da TDM, uma vez que apresentam os valores tanto os de AIC como os de SBC aproximados exceptuando os modelos 1 e 2 em virtude de apresentarem alguns dos seus coeficientes não significativos. No entanto, há que se observar que os modelos 3, 4, 5 e 9 podem ser considerados os que apresentam melhor desempenho uma vez que conseguem explicar uma percentagem elevada da variação do volume de clientes para além de que estes modelos são os que maximizam o logaritmo da verossimilhança. No entanto, para fins de previsão, vai-se utilizar o modelo 3 por apresentar valores de AIC e de SBC baixos e um R^2 elevado comparativamente aos modelos 4, 5 e 9.

4.6 Previsão

A previsão constitui uma parte importante da análise econométrica, para alguns a mais importante (gujarati, 2006).

A partir dos modelos identificados, usou-se o modelo 3 da tabela 4-3, como o melhor e procedeu-se a previsão do volume de clientes dos serviços telefónicos pós-pago residenciais da TDM, SARL, na província de Maputo, para os períodos de Maio a Dezembro de 2008, onde são apresentados os intervalos de confiança para esse período. Com o advento de softwares as previsões foram feitas com recurso ao software estatístico SPSS. Os resultados constam na tabela 4-4.

Tabela 4-4. Previsão do volume de clientes nos serviços telefónicos pós-pago profissionais das TDM, SARL segundo o processo AR(7) para o período de Maio a Dezembro de 2008.

Meses(2008)	volume previsto de clientes	Limite Inferior	Limite Superior
Maio	10506	10393	10619
Junho	10502	1039	10616
Julho	10511	10398	10624
Agosto	10544	10431	10657
Setembro	10543	10431	10656
Outubro	10551	10439	10664
Novembro	10665	10552	10778
Dezembro	10708	10596	10821

Fonte: Adaptação da autora

Neste período, o volume de clientes apresentará uma tendência crescente, passando de aproximadamente 10500 em Maio de 2008 para cerca de 10700 em Dezembro do mesmo ano.

É importante referir que estas previsões dentro da amostra se aproximam aos valores observados para o volume de clientes nos serviços pós-pago profissionais das TDM SARL, na província de Maputo, mostrando desta forma a aplicabilidade dos modelos ARIMA para a previsão do volume de clientes dos serviços telefónicos pós-pago profissionais das TDM, SARL. Os valores observados e previstos apresentam-se na tabela 4-5.

Tabela 4-5: Valores observados e previstos do volume de clientes no período de Maio a Agosto/08 .

Meses(2008)	Valor observado	Valor previsto
Maio	10547	10506
Junho	10633	10502
Julho	10693	10511
Agosto	10759	10544

Fonte: Adaptação da autora

5 CONCLUSÕES E RECOMENDAÇÕES

Neste trabalho fez-se uma análise do comportamento do volume de clientes dos serviços telefónicos pós-pago profissionais das TDM, SARL, com o objectivo de mostrar-se à viabilidade teórica e prática do modelo ARIMA na previsão do volume de clientes, tendo-se concluído o seguinte:

A série do volume de clientes, em estudo, apresentou um comportamento não estacionário em nível, sugerindo que sua média e variância variam ao longo do tempo. Mas, tomadas as primeiras diferenças, a série tornou-se estacionária, cumprindo-se assim, os requisitos para a estimação de modelos de previsão aplicando a metodologia de Box-Jenkins.

Dos modelos estimados, o ARIMA(7,1,0), modelo 3 da tabela 4-3, pode ser aplicado para modelar e prever o volume de clientes.

As previsões para o período de 9/2008 a 12/2008 são: 10543, 10551, 10665 e 10708 com erros de estimação de 225, 225, 226 e 225, respectivamente.

5.1 Recomendações

As conclusões produzidas por este trabalho, sugerem a aplicabilidade dos modelos ARIMA para o estudo e previsão do volume de clientes nos serviços das TDM, SARL, por isso, recomenda-se que se faça um estudo a nível nacional. Recomenda-se ainda, que em futuras pesquisas sobre o volume de clientes, sejam testados os modelos ARCH (Modelos Auto Regressivos de Heterocedasticidade condicional) para captar a volatilidade da série.

5.2 Limitações

A limitação constatada na realização deste trabalho foi a de não poder se extrapolar os resultados obtidos para todo o país, em virtude de se tratar de dados referentes apenas à província de Maputo.

6 BIBLIOGRAFIA

- ARNOLD, J. R. Tony. **Administração de Materiais: uma introdução**, 1999. São Paulo:Atlas.
- Box, G.E.P, Jenkins, G.M, *Time Séries Analysis Forecasting and Control*, 1976. Sam Francisco, Holden.
- BRYSON JR, A. E.; HO, Y. C. **Applied Optimal Control**, 1975. Washington: Hemisphere Publishing Company Corporation.
- CAIADO, Jorge, O Desenvolvimento de Métodos Estatísticos para Análise de Séries Temporais Econômicas e Financeiras, 2004. Revista InforBanca 143.
- CHATTERJEE, S., BERTRAM, P. **Regression analysis by example**, 1991. New York: John Wiley & Sons, Inc.
- CORRÊA, Henrique L.; GIANESI, Irineu G. N.; CAON, Mauro. Planejamento, programação e controle da produção : MRP II/ERP (2000): conceitos, uso e implantação. 3.ed. São Paulo: Atlas,
- DAVIS, Mark M.; AQUILANO, Nicholas J.; CHASE, Richard B. **Fundamentos da administração da produção**, 2001. 3.ed. Porto Alegre : Bookman Editora
- Ehlers, Ricardo Sandes. Análise de Séries Temporárias. 2003.2ª Edição. Brasil
- FONSECA, Jairo S.; MARTINS, Gilberto A.; TOLEDO, Geraldo L. *Estatística Aplicada*, 1995. São Paulo: Atlas.
- GUJARATI, D. N. *Econometria Básica*, 2006 São Paulo: Makron Books.
- LEVINE, David M.; BERENSON, Mark L. STEPHAN, David. **Estatística: Teoria e Aplicações**, 2000. Rio de Janeiro
- MAKRIDAKIS, Spyros G.; WHEELWRIGHT, Steven C.; McGEE, Victor E. *Fundamentals of Quantitative Forecasting*, 1983. New York: John Wiley & Sons.

MAKRIDAKIS, Spyros G.; WHEELWRIGHT, Steven C; HYNDMAN, Rob J. *Forecasting: methods and applications*, 1998, 3.ed. New York: John Wiley & Sons.

MONTGOMERY, Douglas C., RUNGER, George C. **Estatística Aplicada e Probabilidade para Engenheiros**, 2003. 2.ed. Rio de Janeiro. LTC – Livros Técnicos e Científicos Editora S.A.,

MENTZER, John T.; BIENSTOCK, Carol C. **Sales Forecasting Management**, 1998. Thousand Oaks: Sage

MILONE, Giuseppe; ANGELINI, Flávio. **Estatística Aplicada**. 1995 São Paulo: Atlas

MOREIRA, Daniel Augusto. *Administração da produção e operações*, 1996 2.ed. São Paulo: Pioneira.

MORETTIN, Pedro Alberto; TOLOI, Clélia M. *Séries Temporais*, 1987 2.ed. São Paulo: Atual.

MORETTIN, Pedro Alberto. *Econometria Financeira; Um Curso em Séries Temporais Financeiras*, 2004 Departamento de Estatística, Instituto de Matemática e Estatística, USP.

MORETTIN, Pedro Alberto; TOLOI, Clélia M. *Previsão de Séries Temporais*, 1995. São Paulo: Actual Editora.

MORETTIN, Pedro Alberto; TOLOI, Clélia M. *Análise de séries temporais*, 2004. São Paulo: Edgard Blucher.

MORETTIN, Pedro Alberto; TOLOI, Clélia M. *Métodos Quantitativos; Séries Temporais. Previsão de Séries Temporais*, 1986. São Paulo: Editora Atual.

OLIVEIRA, Maria Cristina Neves. *Métodos de Estimação de Parâmetros em Modelos Geoestatísticos com diferentes estruturas de covariâncias*, 2003. Escola Superior de Agronomia “Luiz de Queiroz” – Universidade de São Paulo, Piracicaba.

SALLES, André de Assis. *Estimação do risco sistemático – uma abordagem Bayesiana*. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO. 1996. *Anais...* São Paulo

SÁFADI, Thelma. Uso de séries temporais na análise de vazão de água na represa de FURNAS, 2004, Disponível em: www.editora.ufla.br/revista/28_1/art19.pdf.

STENGEL, Robert F. **Stochastic Optimal Control: Theory and Application**, 1986. 1.ed. New York: John Wiley & Sons Inc

STOCK, J.H. e M.W. WATSON. **Econometria**, 2004. Addison-Wesley

TDM, SARL, Direcção de Marketing: 2008. Perfil da empresa e Catálogo de produtos e serviços

TUBINO, Dalvio F. **Manual de Planeamento e Controlo da Produção**. 2000. 2.ed. São Paulo: Atlas S.A.

ANEXOS

ANEXO A: Dados do volume de clientes dos serviços telefónicos pós-pago residenciais

Tabela 1A: Volume de clientes dos serviços telefónicos pós-pago profissionais de 01/2001 a 08/2008

Meses	2001	2002	2003	2004	2005	2006	2007	2008
Jan-01	10382	10453	9845	9436	9182	9335	9820	10463
Feb-01	10387	10340	9785	9361	9165	9277	9795	10479
Mar-01	10410	10342	9772	9413	9145	9335	9902	10521
Apr-01	10573	10280	9770	9304	9131	9323	9926	10504
May-01	10547	10250	9804	9339	9144	9371	10067	10547
Jun-01	10576	10222	9772	9251	9136	9386	10219	10633
Jul-01	10486	10191	9668	9290	9103	9451	10223	10693
Aug-01	10625	10152	9627	9246	9123	9631	10264	10759
Sep-01	10482	10121	9609	9315	9200	9579	10302	
Oct-01	10518	10066	9623	9304	9164	9576	10401	
Nov-01	10556	9990	9617	9275	9250	9637	10425	
Dec-01	10557	9889	9535	9195	9302	9693	10454	

Fonte: TDM, SARL

Tabela 2A: Estatísticas descritivas do volume de clientes dos serviços telefónicos pós-pago

Estatísticas	
Mean	9854.207
Median	9790.000
Maximum	10759.00
Minimum	9103.000
Std. Dev.	518.2165
Skewness	0.099158
Kurtosis	1.530903
Jarque-Bera	8.424035
Probability	0.014816

Fonte: Adaptação da autora

Figura 1A – Gráfico de Auto correlação da Série do volume de clientes nos serviços telefónicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8.(em nível)

Included observations: 92

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	
*****	*****	1	0.970	0.970	89.498
*****	. .	2	0.941	-0.016	174.55
*****	* .	3	0.907	-0.091	254.46
*****	. .	4	0.871	-0.055	328.97
*****	* .	5	0.830	-0.088	397.50
*****	* .	6	0.785	-0.110	459.42
*****	. .	7	0.739	-0.023	514.96
*****	* .	8	0.685	-0.146	563.32
*****	. .	9	0.632	-0.028	604.92

Fonte: Adaptação da autora

Figura 2A – Gráfico de Auto correlação da Série do volume de clientes nos serviços telefónicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8.(em primeiras diferenças)

Date: 11/13/08 Time: 11:53
 Sample: 2001:01 2008:08
 Included observations: 91

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
. *	. *	1	0.130	0.130	1.6002	0.206
. ***	. ***	2	0.368	0.357	14.498	0.001
. *	. *	3	0.108	0.036	15.619	0.001
. ***	. **	4	0.389	0.288	30.352	0.000
. **	. **	5	0.256	0.207	36.822	0.000
. **	. .	6	0.220	-0.002	41.658	0.000
. ***	. **	7	0.336	0.239	53.011	0.000
. **	. .	8	0.199	0.033	57.049	0.000
. **	. *	9	0.208	-0.077	61.534	0.000
. **	. .	10	0.246	0.156	67.864	0.000
. **	. .	11	0.226	-0.003	73.278	0.000
. **	. .	12	0.233	-0.002	79.104	0.000
. *	. .	13	0.108	-0.030	80.368	0.000
. **	. *	14	0.289	0.074	89.532	0.000
. *	. .	15	0.121	-0.044	91.172	0.000
. *	. *	16	0.164	-0.067	94.193	0.000
. *	. .	17	0.127	0.012	96.024	0.000
. **	. .	18	0.203	0.023	100.79	0.000

Fonte: Adaptação da autora

ANEXO B: Estimação de Modelos

Tabela 1B. Modelo AR (7) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8

Dependent Variable: DCLIENTES
Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000664	0.003389	0.195821	0.8453
AR(2)	0.201775	0.101773	1.982605	0.0509
AR(4)	0.265524	0.099885	2.658298	0.0095
AR(5)	0.129203	0.096432	1.339836	0.1841
AR(7)	0.222564	0.099288	2.241593	0.0278
R-squared	0.304713	Mean dependent var		0.000149
Adjusted R-squared	0.269509	S.D. dependent var		0.006507
S.E. of regression	0.005561	Akaike info criterion		-7.488368
Sum squared resid	0.002443	Schwarz criterion		-7.343677
Log likelihood	319.5115	F-statistic		8.655542
Durbin-Watson stat	1.801945	Prob(F-statistic)		0.000008

Fonte: Adaptação da autora

Tabela 2B. Modelo AR (7) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8

Dependent Variable: DCLIENTES
Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(2)	0.200167	0.101657	1.969037	0.0524
AR(4)	0.274989	0.099505	2.763575	0.0071
AR(7)	0.269661	0.092828	2.904959	0.0047
R-squared	0.288491	Mean dependent var		0.000149
Adjusted R-squared	0.270923	S.D. dependent var		0.006507
S.E. of regression	0.005556	Akaike info criterion		-7.512924
Sum squared resid	0.002500	Schwarz criterion		-7.426110
Log likelihood	318.5428	Durbin-Watson stat		1.821148

Fonte: Adaptação da autora

Tabela 3B. Modelo AR (7) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8

Dependent Variable: DCLIENTES
Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(4)	0.345857	0.094381	3.664463	0.0004
AR(7)	0.314365	0.091574	3.432905	0.0009
R-squared	0.254435	Mean dependent var		0.000149
Adjusted R-squared	0.245342	S.D. dependent var		0.006507
S.E. of regression	0.005652	Akaike info criterion		-7.489979
Sum squared resid	0.002620	Schwarz criterion		-7.432102
Log likelihood	316.5791	Durbin-Watson stat		1.820435

Fonte: Adaptação da autora

Tabela 4B. Modelo AR (7) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8

Dependent Variable: DCLIENTES
Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(2)	0.301782	0.098536	3.062672	0.0030
AR(7)	0.270404	0.096511	2.801784	0.0063
R-squared	0.221405	Mean dependent var		0.000149
Adjusted R-squared	0.211910	S.D. dependent var		0.006507
S.E. of regression	0.005776	Akaike info criterion		-7.446630
Sum squared resid	0.002736	Schwarz criterion		-7.388753
Log likelihood	314.7585	Durbin-Watson stat		1.735111

Fonte: Adaptação da autora

Tabela 5B. Modelo AR (4) do volume de clientes nos serviços telefônicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8

Dependent Variable: DCLIENTES
Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(2)	0.266303	0.099668	2.671894	0.0090
AR(4)	0.305656	0.100661	3.036476	0.0032
R-squared	0.230222	Mean dependent var		0.000229
Adjusted R-squared	0.221165	S.D. dependent var		0.006617
S.E. of regression	0.005839	Akaike info criterion		-7.425674
Sum squared resid	0.002898	Schwarz criterion		-7.368986
Log likelihood	325.0168	Durbin-Watson stat		1.995342

Fonte: Adaptação da autora

Tabela 6B. Modelo AR (2) do volume de clientes nos serviços telefónicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8

Dependent Variable: DCLIENTES
Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(2)	0.375651	0.099726	3.766839	0.0003
R-squared	0.136224	Mean dependent var		0.000371
Adjusted R-squared	0.136224	S.D. dependent var		0.006746
S.E. of regression	0.006270	Akaike info criterion		-7.294957
Sum squared resid	0.003459	Schwarz criterion		-7.266995
Log likelihood	325.6256	Durbin-Watson stat		1.805698

Fonte: Adaptação da autora

Tabela 7B. Modelo AR (4) do volume de clientes nos serviços telefónicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8

Dependent Variable: DCLIENTES
Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(4)	0.402809	0.097157	4.145946	0.0001
R-squared	0.165569	Mean dependent var		0.000229
Adjusted R-squared	0.165569	S.D. dependent var		0.006617
S.E. of regression	0.006044	Akaike info criterion		-7.368015
Sum squared resid	0.003142	Schwarz criterion		-7.339671
Log likelihood	321.5087	Durbin-Watson stat		1.960622

Fonte: Adaptação da autora

Tabela 8B. Modelo AR (7) do volume de clientes nos serviços telefónicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8

Dependent Variable: DCLIENTES
Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(7)	0.348279	0.097688	3.565198	0.0006
R-squared	0.132341	Mean dependent var		0.000149
Adjusted R-squared	0.132341	S.D. dependent var		0.006507
S.E. of regression	0.006061	Akaike info criterion		-7.362132
Sum squared resid	0.003049	Schwarz criterion		-7.333194
Log likelihood	310.2096	Durbin-Watson stat		1.710627

Fonte: Adaptação da autora



Tabela 9B. Modelo ARIMA (4, 1, 2) do volume de clientes nos serviços telefónicos pós-pago profissionais das TDM, SARL, 2001-1 a 2008-8

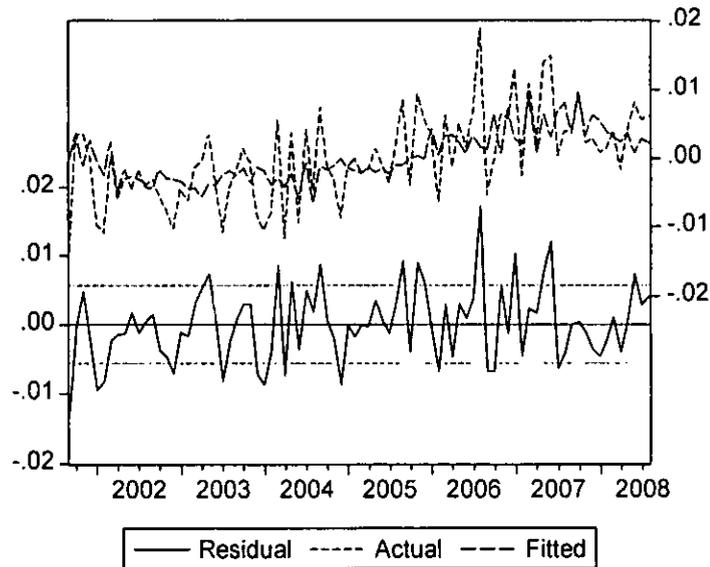
Dependent Variable: DCLIENTES
Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(4)	0.373804	0.100990	3.701413	0.0004
MA(2)	0.295017	0.106930	2.758979	0.0071
R-squared	0.224309	Mean dependent var		0.000229
Adjusted R-squared	0.215183	S.D. dependent var		0.006617
S.E. of regression	0.005862	Akaike info criterion		-7.418022
Sum squared resid	0.002921	Schwarz criterion		-7.361334
Log likelihood	324.6839	Durbin-Watson stat		1.919594

Fonte: Adaptação da autora

ANEXO C— Análise Residual dos modelos estimados

Figura 1C: Gráfico dos resíduos do modelo AR (7)



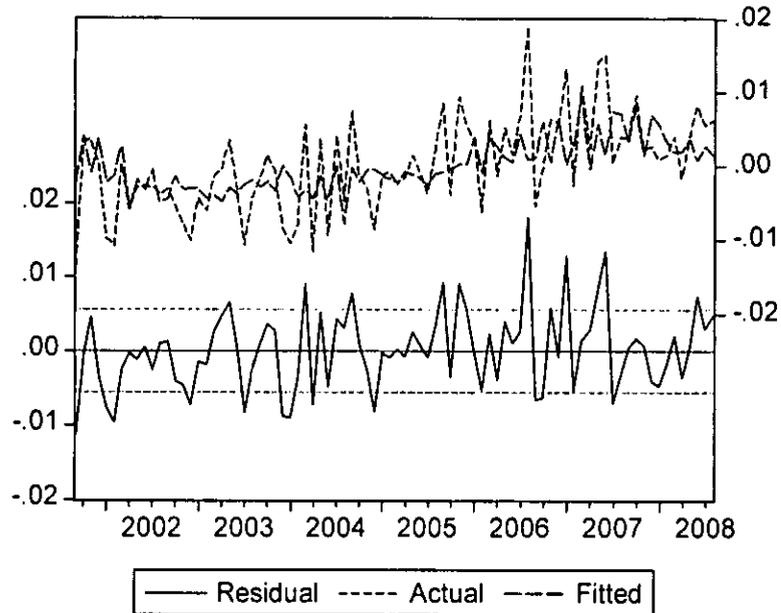
Fonte: Adaptação da autora

Figura 2C: Correlograma da FAC e FACP dos resíduos do modelo AR (7)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
. .	. .	1 0.059	0.059	0.3007	0.583
. .	. .	2 -0.034	-0.037	0.4002	0.819
.* .	.* .	3 -0.117	-0.113	1.6140	0.656
. .	. .	4 0.012	0.024	1.6260	0.804
. * .	. * .	5 0.090	0.082	2.3679	0.796
. .	. .	6 0.000	-0.023	2.3679	0.883
.* .	.* .	7 -0.079	-0.071	2.9506	0.890
. .	. .	8 -0.015	0.013	2.9723	0.936
. .	. .	9 -0.030	-0.039	3.0618	0.962
. * .	. * .	10 0.147	0.131	5.1788	0.879
. .	. .	11 0.016	0.001	5.2029	0.921
. * .	. * .	12 0.073	0.085	5.7310	0.929
. .	. .	13 -0.049	-0.031	5.9750	0.947
. * .	. * .	14 0.169	0.188	8.9394	0.835
. .	. .	15 0.015	-0.019	8.9640	0.879
. * .	. * .	16 0.069	0.077	9.4752	0.893
. .	. .	17 0.004	0.039	9.4771	0.924
. * .	. * .	18 0.148	0.177	11.878	0.853

Fonte: Adaptação da autora

Figura 3C: Gráfico dos resíduos do modelo AR (7)



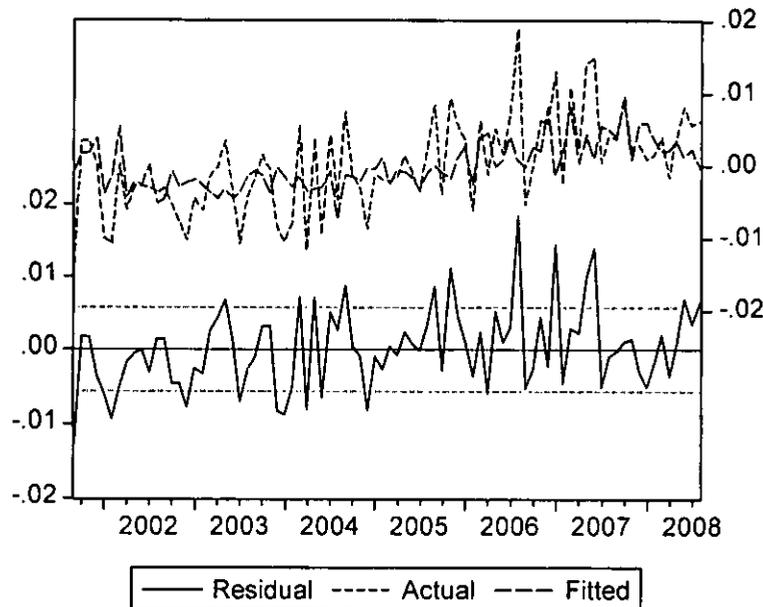
Fonte: Adaptação da autora

Figura 4C: Correlograma da FAC e FACP dos resíduos do modelo AR(7)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
. .	. .	1 0.059	0.059	0.2980	0.585
. .	. .	2 -0.037	-0.041	0.4201	0.811
* .	* .	3 -0.120	-0.116	1.7050	0.636
. .	. .	4 0.000	0.012	1.7050	0.790
. **	. **	5 0.230	0.225	6.5251	0.258
. .	. .	6 0.005	-0.037	6.5270	0.367
* .	* .	7 -0.093	-0.087	7.3411	0.394
. .	. .	8 -0.043	0.021	7.5179	0.482
. .	. .	9 0.018	0.018	7.5500	0.580
. *	. *	10 0.172	0.106	10.454	0.402
. .	. .	11 0.016	0.001	10.479	0.488
. *	. *	12 0.094	0.148	11.371	0.497
. .	. .	13 -0.024	-0.004	11.429	0.575
. *	. *	14 0.167	0.175	14.309	0.427
. .	. .	15 0.011	-0.051	14.321	0.501
. *	. *	16 0.075	0.102	14.924	0.530
. .	. .	17 0.022	0.021	14.975	0.597
. *	. *	18 0.137	0.182	17.028	0.521
. .	. .	19 -0.026	-0.102	17.105	0.583
** .	** .	20 -0.205	-0.214	21.855	0.348
. .	. .	21 -0.043	0.001	22.066	0.396

Fonte: Adaptação da autora

Figura 5C: Gráfico dos resíduos do modelo AR (7)



Fonte: Adaptação da autora

Figura 6C: Correlograma da FAC e FACP dos resíduos do modelo AR (7)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
. .	. .	1	0.049	0.049	0.2117	0.645
. *	. *	2	0.136	0.134	1.8362	0.399
* .	* .	3	-0.070	-0.084	2.2679	0.519
. .	. .	4	-0.014	-0.025	2.2847	0.684
. *	. **	5	0.186	0.215	5.4468	0.364
. .	. .	6	0.042	0.022	5.6128	0.468
* .	* .	7	-0.080	-0.158	6.2133	0.515
. .	. .	8	0.006	0.048	6.2171	0.623
. .	. *	9	0.036	0.101	6.3439	0.705
. *	. *	10	0.195	0.129	10.051	0.436
. .	. .	11	0.012	-0.055	10.065	0.525
. *	. *	12	0.145	0.168	12.173	0.432
. .	. .	13	0.002	0.039	12.173	0.513
. *	. *	14	0.195	0.130	16.094	0.308
. .	. .	15	0.022	-0.050	16.144	0.373
. *	. *	16	0.139	0.140	18.193	0.313
. .	. .	17	0.015	0.026	18.217	0.375
. *	. *	18	0.124	0.092	19.910	0.338
. .	. .	19	-0.031	-0.097	20.018	0.394
* .	** .	20	-0.154	-0.220	22.695	0.304
. .	. .	21	-0.043	-0.002	22.904	0.349

Fonte: Adaptação da autora