

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**

**Departamento de Estatística.**

---

Previsão de séries de Impostos do Município de Belo Horizonte

---



Belo Horizonte - MG

Previsão de séries de Impostos do Município de Belo Horizonte

***TÚLIO MARCOS ALVES AZEREDO***

**Orientadora:** Prof. Marcelo Azevedo Costa

Projeto de monografia apresentado ao Programa de Pós-Graduação do Departamento de Estatística do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais.

UFMG

**Julho de 2010**

## **AGRADECIMENTOS**

Ao meu orientador, professor Marcelo, agradeço pela atenção, esmero, cuidado, paciência, disponibilidade, dispensando incomparável contribuição na realização deste trabalho.

A todos os professores do Curso de Especialização agradeço pelo empenho e competência no exercício da atividade docente.

A minha esposa Maria Cristina, meus filhos: Túlio e Pedro, agradeço pelo incentivo e compreensão, em especial, nos momentos em que estava presente/ausente, estudando e produzindo este trabalho.

## RESUMO

Constitucionalmente, ao poder público em todas as suas esferas foram atribuídas competências ou responsabilidades que fazem parte da essência da sua atividade e que passam pela criação e manutenção de serviços de saúde, educação e segurança, dentre outros.

A tarefa do poder público é administrar a “coisa pública”, de modo a garantir que o cidadão viva em um ambiente estável que lhe proporcione o mínimo de bem estar - no sentido do atendimento de suas necessidades básicas individuais e coletivas. Para esta tarefa é essencial a arrecadação de tributos, sem os quais não seria possível custear os serviços públicos ofertados. Não por acaso, a atividade de arrecadação tributária também foi delineada em nossa Constituição Republicana.

Diante do dilema dado pela demanda por serviços públicos e o seu necessário custeamento, vislumbramos a necessidade de um planejamento consistente, especificamente no que concerne ao *quantum* a ser cobrado de cada cidadão, a fim de que o dimensionamento das despesas seja balizado pela perspectiva das eventuais receitas de impostos e a tarefa constitucional seja cumprida.

Neste sentido, o presente estudo propõe encontrar e testar modelos de previsão que dimensionem, de forma eficiente, as receitas próprias do município de Belo Horizonte. Para tanto, foram testados modelos estatísticos de previsão, com especial ênfase no Modelo de Regressão.

**Palavras-chave:** Receitas Próprias Municipais, Modelos estatísticos de previsão, Regressão, Alisamento Exponencial e Modelos ARIMA.

## **ABSTRACT**

By the Brazilian constitution, the government is responsible for the development and maintenance of health services, education and security of their citizens, among others. In order to ensure that the citizens live in a stable environment that meets their basic needs, the government collects taxes, which funds public services. Therefore, the activity of tax collection has also been outlined in our Republican Constitution.

Faced with the dilemma of the demand for public services versus their cost, we explore the need for a more consistent planning, specifically with regard to the amount of taxes to be charged for every citizen. In this sense, this study explores the ability of the use of regression models to predict efficiently own revenues of the municipality of Belo Horizonte.

**Keywords:** Own Revenues Municipal, Statistical models for forecasting, regression, exponential smoothing and ARIMA models.

## LISTA DE FIGURAS

|           |  |    |
|-----------|--|----|
| Figura 1  | Gráfico da série Imposto Predial.....  | 40 |
| Figura 2  | Gráfico da série Imposto Territorial Urbano.....   | 40 |
| Figura 3  | Gráfico da série Imposto sobre Transmissão de Bens Imóveis.....  | 40 |
| Figura 4  | Gráfico da série ISS retido na fonte.....  | 41 |
| Figura 5  | Gráfico da série ISS próprio.....  | 41 |
| Figura 6  | Gráfico da série ISS – Autônomos e Liberais.....   | 42 |
| Figura 7  | Box-plot mensal de todas as séries estudadas.....  | 43 |
| Figura 8  | Box-plot anual de todas as séries estudadas.....   | 44 |
| Figura 9  | Gráfico da série Log10 Imposto Predial.....  | 45 |
| Figura 10 | Gráfico de probabilidade e histograma da série Log10 Imposto Predial.....  | 46 |
| Figura 11 | Resíduos do ajuste de Regressão para a série Log10 do Imposto Predial.....   | 48 |
| Figura 12 | Teste de normalidade para os Resíduos do ajuste de regressão para a série Log10 do Imposto Predial.....            | 48 |
| Figura 13 | FAC e FACP para os Resíduos do ajuste de regressão para a série Log10 do Imposto Predial.....                      | 49 |
| Figura 14 | Resíduos do ajuste de regressão para a série Log10 do Imposto Predial.....   | 51 |
| Figura 15 | Teste de normalidade para os resíduos do ajuste de regressão da série Log10 do Imposto Predial.....                | 51 |
| Figura 16 | Gráfico da série Log10 do Imposto Territorial Urbano.....  | 53 |
| Figura 17 | Gráfico de probabilidade e histograma série Log10 do Imposto Territorial Urbano.....                               | 54 |
| Figura 18 | Resíduos do ajuste de regressão para a série Log10 do Imposto Territorial Urbano.....                              | 55 |
| Figura 19 | Teste de normalidade para os resíduos do ajuste de regressão para a série Log10 do Imposto Territorial Urbano..... | 56 |
| Figura 20 | FAC e FACP para os resíduos do ajuste de regressão da série Log10 do Imposto Territorial Urbano.....               | 56 |
| Figura 21 | Resíduos do ajuste de regressão para série Log10 do Imposto Territorial Urbano.....                                | 59 |
| Figura 22 | Teste de normalidade para os resíduos do ajuste de regressão para a série Log10 do Imposto Territorial Urbano..... | 59 |
| Figura 23 | Gráfico da série Log10 ITBI.....   | 61 |
| Figura 24 | Gráfico de probabilidade e histograma da série Log10 ITBI.....   | 61 |
| Figura 25 | Alisamento Holt Winters para a série Log10 ITBI.....   | 62 |
| Figura 26 | Resíduos do alisamento exponencial Holt Winters para a série Log10 ITBI.....                                       | 63 |
| Figura 27 | Teste de normalidade para os resíduos da série Log10 ITBI.....   | 63 |

|           |   |    |
|-----------|---|----|
| Figura 28 | FAC e FACP para os resíduos do ajuste por alisamento Holt Winters para série Log10 ITBI.....              | 64 |
| Figura 29 | Gráfico da série Log10 ISS.....   | 65 |
| Figura 30 | Gráfico de probabilidade e histograma da série Log10 ISS.....   | 66 |
| Figura 31 | Alisamento exponencial Holt Winters para a série Log10 ISS.....   | 66 |
| Figura 32 | Resíduos do alisamento exponencial Holt Winters para a série Log10 ISS.....                               | 67 |
| Figura 33 | FAC e FACP para os resíduos do ajuste por alisamento exponencial Holt Winters para a série Log10 ISS..... | 67 |
| Figura 34 | Gráfico da série Log10 ISS Autônomos e Liberais.....  | 69 |
| Figura 35 | Gráfico de probabilidade e histograma série Log10 ISS Autônomos e Liberais                                | 69 |
| Figura 36 | Resíduos do ajuste de regressão para a série Log10 ISS Autônomos e Liberais.....                          | 71 |
| Figura 37 | Teste de normalidade para os resíduos do ajuste de regressão para o Log10 ISS Autônomos e Liberais.....   | 72 |
| Figura 38 | FAC e FACP para os resíduos do ajuste de regressão para a série Log10 Autônomos e Liberais.....           | 72 |
| Figura 39 | Resíduos do ajuste de regressão para o Log10 ISS Autônomos e Liberais.....                                | 75 |
| Figura 40 | Teste de normalidade para os resíduos do ajuste de regressão para o Log10 ISS Autônomos e Liberais.....   | 75 |

## LISTA DE TABELAS E QUADROS

|           |  |    |
|-----------|--|----|
| Quadro 1  | Estatísticas descritivas das séries temporais analisadas.....  | 44 |
| Tabela 1  | Estimativas dos parâmetros de regressão para a série Log10 Imposto Predial....                               | 47 |
| Tabela 2  | Estimativas dos parâmetros de regressão para a série Log10 Imposto Predial....                               | 50 |
| Tabela 3  | Previsões (1 passo a frente) maio a setembro de 2009 – Imposto predial.....                                  | 52 |
| Tabela 4  | Estimativas dos parâmetros de regressão para a série Log10 Imposto Territorial                               | 55 |
| Tabela 5  | Estimativas dos parâmetros de regressão para a série Log10 Imposto Territorial                               | 58 |
| Tabela 6  | Previsões (1 passo a frente) maio a setembro de 2009 – Imposto Territorial.....                              | 60 |
| Tabela 7  | Previsões (1 passo a frente) maio a setembro de 2009 – Imposto sobre Transmissão de Bens Imóveis (ITBI)..... | 64 |
| Tabela 8  | Previsões (1 passo a frente) maio a setembro de 2009 – ISS.....  | 68 |
| Tabela 9  | Estimativas dos parâmetros de regressão para a série Log10 ISS autônomos e Liberais.....                     | 70 |
| Tabela 10 | Estimativas dos parâmetros de regressão para a série Log10 ISS autônomos e Liberais.....                     | 74 |
| Tabela 11 | Previsões (1 passo a frente) maio a setembro de 2009 – Log 10 ISS Autônomos e Liberais.....                  | 76 |

## SUMÁRIO

|   |    |
|---|----|
| 1 – INTRODUÇÃO.....   | 11 |
| 2 – METODOLOGIA.....  | 13 |
| 2.1 - PROCESSOS ESTOCÁSTICOS E SÉRIES TEMPORAIS.....                        | 13 |
| 2.1.1 – Estacionariedade.....   | 14 |
| 2.1.2 – Diferenciação.....  | 15 |
| 2.1.3 - Transformação Box-Cox.....  | 15 |
| 2.1.4 - Análise de covariância e correlação.....                            | 16 |
| 2.1.5 - Ruído branco.....   | 19 |
| 2.2 - MODELAGEM DE SÉRIES TEMPORAIS.....                                    | 20 |
| 2.2.1 - Regressão em séries temporais.....                                  | 20 |
| 2.3 - MÉTODOS SIMPLES DE PREVISÃO DE SÉRIES TEMPORAIS.....                  | 26 |
| 2.3.1 - Alisamento Exponencial Duplo.....                                   | 26 |
| 2.3.2 - Alisamento Exponencial Holt-Winters.....                            | 27 |
| 2.3.3 - Comparação entre modelos de Alisamento Exponencial.....             | 28 |
| 2.4 - MÉTODOS AVANÇADOS DE PREVISÃO DE SÉRIES TEMPORAIS.....                | 29 |
| 2.4.1 - Modelos Autoregressivo e de Médias Móveis.....                      | 29 |
| 2.4.2 - Metodologia de Box e Jenkins.....                                   | 31 |
| 3 - ESCOLHA E AJUSTE DE MODELOS.....  | 39 |
| 3.1 - DADOS (análise preliminar).....                                       | 39 |
| 3.2 - ANÁLISE DESCRITIVA DOS DADOS - IMPOSTO PREDIAL.....                   | 45 |
| 3.2.1 - Ajuste do modelo de regressão.....                                  | 46 |
| 3.2.2 - Ajuste do modelo de regressão com inserção do erro de previsão..... | 49 |
| 3.2.3 - Previsão usando o modelo de regressão ajustado.....                 | 52 |
| 3.3 - ANÁLISE DESCRITIVA DOS DADOS - IMPOSTO TERRITORIAL URBANO.....        | 53 |
| 3.3.1 - Ajuste do modelo de regressão.....                                  | 54 |
| 3.3.2 - Ajuste do modelo de regressão com inserção do erro de previsão..... | 57 |
| 3.3.3 - Previsão usando o modelo de regressão ajustado.....                 | 60 |

|  |    |
|--|----|
| 3.4 - ANÁLISE DESCRITIVA DOS DADOS - IMPOSTO SOBRE TRANSMISSÃO DE BENS IMÓVEIS (ITBI)..... | 61 |
| 3.4.1 - Ajuste do modelo de Alisamento Exponencial Holt Winters.....                       | 62 |
| 3.4.2 - Previsão usando o modelo ajustado.....   | 64 |
| 3.5 - ANÁLISE DESCRITIVA DOS DADOS – IMPOSTO SOBRE SERVIÇOS (ISS).....                     | 65 |
| 3.5.1 - Ajuste do modelo de Alisamento Exponencial Holt Winters.....                       | 66 |
| 3.5.2 - Previsão usando o modelo ajustado.....   | 68 |
| 3.6 - ANÁLISE DESCRITIVA DOS DADOS – ISS AUTONOMOS E PROFISSIONAIS LIBERAIS.....           | 69 |
| 3.6.1 - Ajuste do modelo de regressão.....   | 70 |
| 3.6.2 - Ajuste do modelo de regressão com inserção do erro de previsão.....                | 73 |
| 3.6.3 - Previsão usando o modelo de regressão ajustado.....                                | 76 |
| 4 – DISCUSSÕES E CONCLUSÃO.....  | 77 |
| 5 – BIBLIOGRAFIA.....  | 80 |

# 1 INTRODUÇÃO

Os municípios brasileiros estão autorizados pela Constituição Federal a organizar e administrar seu próprio sistema tributário, desde que respeitadas as condições estabelecidas em diversos regulamentos que disciplinam o Sistema Tributário Nacional.

Compondo este Sistema, existe uma série de impostos que são de exclusividade dos municípios. Em outras palavras, somente os municípios podem instituir e cobrar determinados impostos, pois estão no campo da competência tributária privativa exclusiva dos municípios, conforme determinação legal.

Para o fiel cumprimento desta tarefa constitucional, os municípios devem preparar seus instrumentos legais e de planejamento para determinar o montante de imposto que esperam arrecadar. Esta tarefa é de fundamental importância, pois afeta diretamente a dinâmica administrativa do município. A partir da previsão da arrecadação é que se projetam as despesas com educação, saúde, transporte, obras, dentre outras. Neste sentido, percebe-se a importância de previsões realistas, apoiadas em dados técnicos, pois tais previsões impactam diretamente no sucesso administrativo de qualquer administração pública municipal.

Para sermos mais específicos, trataremos neste trabalho dos impostos que formam as receitas próprias do município. Ou seja, receitas que dependem do esforço arrecadatório do município para que efetivem sua entrada nos cofres públicos dos valores relativos aos seguintes impostos:

1. Imposto sobre Serviços de Qualquer Natureza – ISSQN, desmembrado pelo tipo de recolhimento em Imposto próprio – recolhido pelo próprio contribuinte, Imposto retido na fonte – recolhido pelo tomador do serviço e o ISSQN pagos pelos profissionais autônomos e liberais.
2. Imposto Predial e Territorial Urbano – IPTU, desmembrado em Imposto Predial e Imposto Territorial;
3. Imposto sobre Transmissão de Bens Imóveis – ITBI;

O IPTU e o ISSQN tiveram o recolhimento desmembrado, respectivamente, em duas e três modalidades, em função das características da incidência e do recolhimento do imposto.

As séries temporais estudadas possuem comportamentos distintos, relacionados e explicados pela natureza do imposto, sua incidência e suas peculiaridades. A análise das séries deverá levar em conta os comportamentos diferenciados de cada série temporal.

A proposta deste trabalho é descrever metodologias apropriadas para modelagem de várias séries temporais, relativas à arrecadação de impostos administrados pelo município de Belo Horizonte. Levando em consideração as modalidades de recolhimento, serão trabalhadas seis séries temporais diferentes.

Para identificar modelos de previsão para cada uma das séries temporais estudadas, descreveremos o perfil do montante arrecadado por imposto, por meio da análise dos dados históricos, a fim de identificar tendências que determinem o montante da arrecadação do município, no que se refere ao esforço próprio de arrecadação.

Para analisar as séries e prever seus valores futuros utilizaremos os seguintes procedimentos: Modelos de Regressão para séries temporais (BOWERMAN e O'CONNEL, 1993), Alisamento Exponencial de Holter Winters (MORETTIN e TOLOI, 2004) e o Modelo SARIMA (BOX e JENKINS, 1976). Estes modelos serão avaliados, levando-se em consideração a adequação à série estudada e a capacidade de previsão que proporciona.

A ferramenta computacional utilizada nesse trabalho foi o Minitab em sua versão número 15.

Este trabalho está organizado da seguinte forma: Na segunda seção apresentamos alguns conceitos elementares do estudo de séries temporais e as metodologias utilizadas para analisar as séries. Na terceira seção, apresentam-se os modelos ajustados e suas respectivas previsões. Na quarta seção, são apresentadas as conclusões.

## 2 METODOLOGIA

Neste capítulo descreveremos o embasamento teórico para estudo de processos estocásticos, no qual se inserem as séries temporais relativas a impostos municipais e propomos modelos para previsão e modelagem das séries.

Quando for necessário, as equações estarão identificadas pelo número da página e a posição que ocupa na mesma página em relação às demais equações. Por exemplo: a equação 3.2 estará localizada na página 3 e será a segunda equação a constar na citada página. Com este procedimento, esperamos facilitar a análise e leitura do texto, facilitando a localização das equações e suas interrelações.

### 2.1 – PROCESSOS ESTOCÁSTICOS E SÉRIES TEMPORAIS

Um processo estocástico é uma série de variáveis aleatórias ordenadas no tempo, assumindo valores no espaço de índices “ $t$ ”, onde este espaço pode ser discreto ou contínuo.

$$\{ Y_t, t = 1, 2, \dots \}$$

Uma série temporal, denotada aqui por “ $Y_t$ ” é um conjunto de observações obtidas ao longo do tempo. Considera-se que a série está sendo observada a partir de um ponto específico no passado, definindo como ponto de origem ou ponto “1”, até o instante “ $n$ ”, *assim, temos:*

$$\{ Y_1, Y_2, \dots, Y_n \}$$

Através do estudo sistemático das propriedades estatísticas e assumindo um modelo ajustado à série, podemos estimar valores futuros, *assim, teremos:*

$$\{ Y_{n+1}, Y_{n+2}, \dots, Y_{n+m} \}$$

### 2.1.1 - Estacionariedade

“O conceito de estacionariedade é a principal idéia que se deve ter para estimar uma série temporal” (BUENO, 2008).

“Uma série é dita estacionária se sua média e sua variância são constantes ao longo do tempo e se o valor da covariância entre dois períodos de tempo depende apenas da distância ou da defasagem (*lag*) entre esse dois períodos de tempo e não do período de tempo específico no qual a covariância é calculada” (SHIKIDA, 2008), indicando um padrão de equilíbrio temporal.

Uma série  $Y_t$  será estritamente estacionária, se a distribuição de probabilidade não muda com o tempo e, diz-se que um processo é fracamente estacionário (ou estacionário de 2ª ordem) se e somente se:

$$E(Y_t) = \mu, \text{ constante para todo } t;$$

$$VAR(Y_t) = \sigma^2 \text{ é finita e constante para todo } t;$$

$$Cov(Y_t, Y_s) = \text{função que depende somente do modulo da diferença } |t - s| .$$

A maioria dos procedimentos de análise de séries temporais supõe que as séries são estacionárias, porém, em situações reais, é improvável encontrar séries perfeitamente estacionárias.

Para resolver o problema da não estacionariedade, aplicam-se dois procedimentos, o primeiro para o ajuste da média e o segundo, para o ajuste da variância. Em relação à média, podem-se realizar operações sucessivas de diferenciação (BOX e JENKINS, 1976) para se obter uma média constante e, quando a série não é constante na variância, pode-se aplicar uma transformação de BOX-COX.

### 2.1.2 - Diferenciação

Aplicar diferenciações à série é uma forma de transformá-la. Neste procedimento o operador de defasagem será aplicado tomando diferenças sucessivas, até tornar a série estacionária.

A primeira diferença é definida por:

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

Onde,  $Y_t$  é a observação da série no tempo  $t$ ;

$Y_{t-1}$  é a observação da série no tempo  $t-1$ .

A segunda diferença é:

$$\begin{aligned}\Delta(Y_t - Y_{t-1}) &= \Delta Y_t - \Delta Y_{t-1} \\ &= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \\ &= Y_t - 2Y_{t-1} + Y_{t-2}\end{aligned}$$

De modo geral, temos que a  $n$ -ésima diferença  $Y_{(t)}$  é dada por:

$$\Delta^d Y_t = \Delta[\Delta^{d-1} Y_t]$$

### 2.1.3 - Transformação Box-Cox

Através do método denominado de “transformação Box-Cox”, podemos obter uma série com variância aproximadamente constante, mais simétrica e próxima da distribuição normal.

$$Y_t^{(\lambda)} = \begin{cases} \frac{Y_t^\lambda - c}{\lambda}, & \text{se } \lambda \neq 0 \\ \log Y_t, & \lambda = 0. \end{cases}$$

O valor de  $\lambda$  e  $c$  são parâmetros a serem estimados.

#### 2.1.4 - Análise de covariância e correlação

As estatísticas que medem a relação entre as observações da mesma série em diferentes instantes de tempo  $Y_t$  e  $Y_{t-k}$  são chamadas de autocovariância e autocorrelação.

##### 2.1.4.1 Função de autocovariância

A função de autocovariância entre observações nos instantes  $t$  e  $t+k$ , para uma série estacionária, é definida da seguinte forma:

$$\gamma_k = \text{Cov}(Y_t, Y_{t+k}) = E[(Y_t - \mu)(Y_{t+k} - \mu)]$$

*Com as propriedades:*

$$\gamma_0 = \text{Var}(Y_t),$$

$$|\gamma_k| \leq \gamma_0; \text{ para todo } k;$$

$$\gamma_{-k} = \gamma_k; \text{ para todo } k.$$

##### 2.1.4.2 Função de autocorrelação

A função de autocorrelação (FAC) descreve a correlação observada entre dois valores da mesma série, em diferentes períodos de tempo.

A função de autocorrelação é utilizada para medir o comprimento e a memória de um processo; ou seja, a extensão para a qual o valor tomado no tempo  $t$  depende daquele tomado no tempo  $t+k$  ou  $t-k$ .

A autocorrelação de defasagem  $k$  entre duas variáveis é dada por:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{Cov}[Y_t, Y_{t+k}]}{\sqrt{\text{Var}(Y_t) \cdot \text{Var}(Y_{t+k})}}, \quad K = 0, \pm 1, \pm 2, \dots$$

onde  $\gamma_k$  corresponde à autocovariância entre  $Y_t$  e  $Y_{t+k}$ .

Assim, temos que:

$$\text{Var}(Y_t) = \text{Var}(Y_{t+k}) = \gamma_0 = \text{variância do processo.}$$

Com :  $\rho_0 = 1$  e,

$$\rho_k = \rho - k$$

A Função de Autocorrelação Parcial (FACP) é utilizada para analisar a correlação entre observações em diferentes instantes de tempo. Mas, neste caso, ao medirmos a correlação entre duas observações seriais  $Y_t$  e  $Y_{t+k+1}$ , primeiro eliminamos a dependência dos termos intermediários  $Y_{t+1}, Y_{t+2}, \dots, Y_{t+k-1}$ . Dessa forma, teremos a autocorrelação parcial, representada por:

$$\text{Cov}(Y_t, Y_{t+k} | Y_{t+1}, \dots, Y_{t+k-1})$$

A função de autocorrelação parcial pode ser obtida, considerando-se um modelo de regressão para um processo estacionário, com média zero.

A variável dependente  $Y_{t+k}$  depende das variáveis  $Y_{t+k-1}, Y_{t+k-2}, \dots$

Assim,

$$Y_{t+k} = \phi_{k+1} Y_{t+k-1} + \phi_{k+2} Y_{t+k-2} + \dots + \phi_{kk} Y_t + a_{t+k} \quad (7.1)$$

Onde  $\phi_{kt}$  - i-ésimo parâmetro da regressão;

$a_{t+k}$  - é o termo do erro não correlacionado com  $Y_{t+k-j}$ , para  $j \geq 1$ .

Multiplicando-se  $Y_{t+k-j}$  por ambos os lados em (7.1) e aplicando a esperança, temos:

$$\gamma_j = \phi_{k1}\gamma_{j-1} + \phi_{k2}\gamma_{j-2} + \dots + \phi_{kk}\gamma_{j-k} .$$

Então,

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \dots + \phi_{kk}\rho_{j-k} .$$

Para  $j = 1, 2, \dots, k$  obtemos o sistema de equações:

$$\rho_1 = \phi_{k1}\rho_0 + \phi_{k2}\rho_1 + \dots + \phi_{kk}\rho_{k-1}$$

$$\rho_2 = \phi_{k1}\rho_1 + \phi_{k2}\rho_0 + \dots + \phi_{kk}\rho_{k-2}$$

$$\rho_k = \phi_{k1}\rho_{k-1} + \phi_{k2}\rho_{k-2} + \dots + \phi_{kk}\rho_0$$

Usando a “regra de Cramer” para  $k = 1, 2, 3, \dots, k$  temos:

$$\phi_{11} = \rho_1 \quad \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}$$

$$\text{De forma geral, temos: } \phi_{kk} = \frac{|P_k^*|}{|P_k|}$$

Onde  $P_k$  é a matriz de autocorrelação e  $P_k^*$  é a matriz  $P_k$  com a última coluna substituída pelo vetor de autocorrelações.

### 2.1.5 - Ruído branco

Uma sequência  $\varepsilon_t$  é considerada um ruído branco se para cada valor da série, tem-se média zero, variância constante e autocorrelação igual a zero. A média zero é conveniência, pois seria possível especificar um ruído branco cuja média fosse diferente de zero. Entretanto, pode-se centrar em zero tal série, sem prejuízo de suas demais propriedades (BUENO, 2008, p.19).

Em suma, uma sequência  $\{\varepsilon_t\}_{t=-\infty}^{\infty}$  de variáveis aleatórias é dita um ruído branco, se:

$$E\{\varepsilon_t\} = 0; \forall t$$

$$E\{\varepsilon_t^2\} = \sigma^2; \forall t$$

$$COV(\varepsilon_t, \varepsilon_{t-j}) = 0; \text{ todo } j \neq 0.$$

A representação de um ruído branco pode ser definida como  $RB(0, \sigma^2)$ . Assim, O processo de  $a_t = 1, 2, \dots, i$  pode ser classificado como ruído branco Gaussiano se:

$$a_t \sim N(0, \sigma^2) \text{ e } cov(a_t, a_{t-k}) = 0, \text{ para qualquer valor positivo de } k.$$

Conforme Bueno (2008, p.19), tal processo é chamado ruído branco porque sua função densidade espectral é horizontal como a luz branca, e o processo provoca alterações na série assim com as ondas eletromagnéticas produzem ruídos na sintonização de um rádio. Portanto, um ruído branco é, ao mesmo tempo, temporalmente homogêneo, estacionário e sem memória. Ou seja, sem dependência temporal.

## 2.2 – MODELAGEM DE SÉRIES TEMPORAIS

Com base na análise descritiva de uma série temporal  $Y_t$ , podemos encontrar, entre os modelos disponíveis, aqueles que melhor explicam a dependência temporal. Neste trabalho, utilizaremos os modelos de regressão em séries temporais, alisamento exponencial e modelos SARIMA de Box e Jenkins.

### 2.2.1 - Regressão em séries temporais

“O objetivo primário da análise de regressão linear é estimar o valor esperado de uma variável aleatória (a variável dependente) dado que o valor de uma variável associada (a variável independente) é conhecido” (KAZMIER, 2007, p. 255).

A equação linear que representa o modelo de regressão linear simples é:

$$Y_t = \beta_0 + \beta_1 x_t + a_t$$

onde:

$Y_t$  é o valor da variável dependente na  $t$ -ésima tentativa, ou observação

$\beta_0$  é o primeiro parâmetro da equação de regressão, ou intercepto, o qual indica o valor de  $y$  quando  $X=0$

$\beta_1$  é o segundo parâmetro da equação de regressão, chamado de coeficiente de regressão, que indica a inclinação da reta de regressão

$X_t$  é o valor especificado da variável independente na  $t$ -ésima tentativa, ou observação

$a_t$  representa o erro, ou seja, a componente aleatória na  $t$ -ésima tentativa

A equação ajustada pelo modelo de regressão linear simples é dada por:

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_t$$

sendo  $t = 1, 2, \dots, n$  e os parâmetros do modelo,  $\beta_0$  e  $\beta_1$ , estimados através do Método de Mínimos Quadrados (KAZMIER, 2007, p. 257).

#### 2.2.1.1 - Adequação dos modelos de regressão:

Para verificar se o modelo de regressão é adequado são utilizadas algumas ou todas as seguintes estatísticas de teste: (a) verificar se as variáveis são significativas através de teste de hipótese para os parâmetros  $\beta_0$  e  $\beta_1$ , (b) analisar o coeficiente AIC, (c) verificar o coeficiente de determinação, (d) analisar os Resíduos.

##### a) Teste de hipótese:

Testa-se a hipótese de significância dos parâmetros, na forma:

$$H_0: \beta_0, \beta_1 = 0;$$

$$H_1: \beta_0, \beta_1 \neq 0;$$

Ao nível de significância de 5%, se o nível descritivo ou  $p$ -valor for superior a 0,05, assume-se que não existe evidência suficiente para rejeitar a hipótese nula. Logo, há uma forte evidência de que o coeficiente em análise ( $\beta_0$  ou  $\beta_1$ ) tenha seu valor, senão igual, muito próximo de zero. Em ocorrendo esta situação o coeficiente deve ser excluído da regressão. Caso contrário, mantêm-se o coeficiente na equação de regressão.

##### b) Coeficiente AIC:

O critério de Informação de Akaike (AIC – *Akaike's Information Criterion*) procura uma solução satisfatória entre o bom ajuste e o princípio da parcimônia (AKAIKE, 1973), ou seja, um modelo com baixo erro e poucas variáveis preditoras. Este critério é definido por:

$$AIC(q) = nLn \sigma_a^2 + 2M$$

Onde  $M$  é o número de parâmetros estimado e  $\sigma_a^2$  é a variância do vetor de resíduos.

c) - Coeficiente de determinação ( $R^2$ ):

O coeficiente de determinação mede, em termos percentuais, a variação total explicada pelo modelo de regressão analisado. O coeficiente de determinação de uma regressão assim se define:

$$R^2 = \frac{SQ\text{ Reg}}{STQ} \quad \text{onde, } 0 \leq R^2 \leq 1$$

A Soma de Quadrados da Regressão ( $SQ\text{ Reg}$ ) representa a soma dos quadrados dos desvios dos valores preditos em relação à sua média e é calculada da seguinte forma:

$$\sum_t (\hat{Y}_t - \bar{Y})^2, \quad t=1, \dots, n$$

A Soma de Quadrados Total ( $STQ$ ) é dada pela variação total da variável resposta em torno de sua média e é calculada da seguinte forma:

$$SQT = \sum_t (Y_t - \bar{Y})^2, \quad t=1, \dots, n$$

d) - Análise dos resíduos:

A análise dos resíduos resultantes da regressão é fundamental para definir se o modelo está bem ajustado. Essa validação é realizada através da análise dos resíduos estimados, dados por:  $\hat{e}_t = y_t - \hat{y}_t$

A variável  $a_t$  representa o erro do modelo e deve satisfazer algumas condições para que o modelo seja válido.

(i) O termo de erro tem média zero -  $\bar{e} = \frac{\sum e_i}{n} = 0$  ;

(ii) O erro é homocedástico, ou seja, variância constante -  $\frac{\sum (e_i - \bar{e})^2}{n-2} = \frac{\sum e_i^2}{n-2} = \frac{SQE}{n-2} = EQM$ ,

Se o modelo estiver adequado, o EQM (Erro Quadrático Médio) é um estimador não tendencioso da variância ( $\sigma^2$ ) do erro ( $a_t$ ).

A validação da suposição de que a variância é constante é feita através da análise gráfica dos resíduos versus valores ajustados. Caso a suposição seja válida, os resíduos apresentarão um bom ajuste, distribuindo-se aleatoriamente em torno da média zero.

(iii) O erro tem distribuição normal  $a_t^{iid} \sim N(0, \sigma^2)$ . Esta suposição é fundamental para que se possa usar a inferência estatística na distribuição dos estimadores e na análise de variância do modelo.

Para a verificação da normalidade do modelo, será testada a hipótese nula de que os resíduos estão distribuídos tal qual uma Distribuição Normal. Se o *p-valor* obtido para o teste de normalidade for maior que o nível de significância assumido de 5%, haverá evidências para rejeitar a hipótese nula, assumindo que os resíduos se aproximam da distribuição normal.

(iiii) Outra importante suposição considera que os erros individuais são independentes entre si;

Na verdade, os resíduos não são variáveis aleatórias independentes, pois eles envolvem os valores estimados ( $\hat{Y}_t$ ) os quais são baseados na mesma equação de regressão.

Para testar a suposição de independência, pode-se aplicar o teste de DURBIN e WATSON (1951):

A estatística do teste é dada por: 
$$d = \frac{\sum_{t=2}^n (\hat{a}_t - \hat{a}_{t-1})^2}{\sum_{t=1}^n \hat{a}_t^2}$$

Onde  $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$ , representam os erros ordenados no tempo.

Para verificação de autocorrelação de primeira ordem, positiva ou negativa, são testadas duas hipóteses:

$H_0$  = os termos de erro não são correlacionados.

$H_1$  = os termos de erro são correlacionados.

Caso:  $d < d_{L,\alpha}$  ou se,  $(4-d) < d_{L,\alpha}$ , rejeita-se  $H_0$

$d > d_{U,\alpha}$  e  $(4-d) > d_{U,\alpha}$ , não rejeita-se  $H_0$

$d_{L,\alpha} \leq d \leq d_{U,\alpha}$  e  $d_{L,\alpha} \leq (4-d) \leq d_{U,\alpha}$  o teste é inconclusivo.

Onde  $d_L$  e  $d_U$  são, respectivamente, os pontos críticos de mínimo e máximo do intervalo de confiança, encontrados na Tabela de DURBIN e WATSON (1951) e o " $\alpha$ " o nível de significância.

Se o modelo de regressão atender a todas as suposições anteriormente descritas poderá ser utilizado para modelar os dados e prever valores futuros.

Caso contrário, se o modelo de regressão não atender a todas as suposições, em particular a que se refere à independência dos resíduos, pode-se ajustar um modelo que inclua correlação de primeira ordem para os resíduos da série estudada (BOWERMAN e O'CONNEL, 1993).

### 2.2.1.2 - Correlação de 1º ordem

A correlação de primeira ordem é dada por:  $a_t = \phi_1 a_{t-1} + \varepsilon_t$

E o modelo de regressão se torna:  $Y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_p x_{tp} + a_t$

No qual  $a_t$  é descrito por um processo autoregressivo.

Considerando que  $\varepsilon_t$  é o termo de erro com média zero, variância constante, independente e atende a suposição de normalidade, supõe-se que  $\varepsilon_t$  será definido por um processo autoregressivo de primeira ordem.

A equação para predição de valor futuro  $Y_{t+r}$  é:

$$\begin{aligned} Y_{t+r} &= \mu_{t+r} + a_{t+r} \\ &= \beta_0 + \beta_1 x_{1,t+r} + \beta_2 x_{2,t+r} + \dots + \beta_p x_{p,t+r} + a_{t+r} \end{aligned}$$

$$\text{onde } a_{t+r} = \phi_1 a_{t+r-1} + \varepsilon_{t+r1}$$

A equação final de predição será, então, definida a partir dos valores estimados dos parâmetros, da seguinte forma:

$$\hat{Y}_{t+r} = \hat{\mu}_{t+r} + \hat{a}_{t+r} = \hat{\mu}_{t+r} + \hat{\phi}_1 \hat{a}_{t+r-1} + \hat{\varepsilon}_{t+r}$$

## 2.3 - MÉTODOS SIMPLES DE PREVISÃO DE SÉRIES TEMPORAIS

Os métodos assim denominados efetuam a previsão do valor futuro da série temporal pela soma ponderada das observações passadas da série (MAKRIDAKIS, 1985), assumindo que os valores extremos da série representam flutuações aleatórias (MORETTIN e TOLOI, 2006). O propósito desses métodos consiste em identificar um padrão suavizado presente nos dados históricos e, então, usar esse padrão para prever valores futuros.

MORETTIN e TOLOI (2006) associam a grande popularidade desses métodos à simplicidade, à eficiência computacional e à razoável previsão obtida. Entre os métodos simples de previsão destacam-se o de *Médias Móveis*, o *Alisamento Exponencial Simples*, o *Alisamento Exponencial Duplo* e o *Alisamento Exponencial de Holt-Winters*.

Para os objetivos deste trabalho, apresentaremos os modelos de alisamento exponencial duplo e o alisamento exponencial Holt-Winters.

### 2.3.1 - Alisamento Exponencial Duplo

O Alisamento Exponencial Duplo, também denominado de Alisamento de *Holter*, consiste em modelar o nível e a tendência apresentada no conjunto de dados. Os valores do nível e da tendência da série no instante “*t*” serão estimados respectivamente por:

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad 0 < \alpha < 1 \quad t = 2, \dots, n$$

$$T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1} \quad 0 < \gamma < 1 \quad t = 2, \dots, n$$

Onde:

$Y_t$  é a série temporal;

$L_t$  é denominado de valor exponencial suavizado;

$T_t$  é a suavização da tendência; e

$\alpha, \gamma$  são os parâmetros de suavização. A previsão para o valor “ $n$ ” passos à frente da série  $Y_t$ ,  $t = 1, 2, \dots, n$ , com base nas observações até o instante de tempo “ $t$ ”, é dada

por:  $\hat{Y}_t(n) = L_t + nT_t$ ,  $\forall r > 0$ ,

### 2.3.1.1 - Estimação dos parâmetros de suavização

Os valores dos parâmetros de suavização  $\alpha, \gamma$  podem ser associados à influência das observações passadas nas previsões futuras da série. Para valores pequenos de  $\alpha, \gamma$  próximos de zero, vale dizer que as previsões dependem das observações passadas e, para valores próximos de 1, as previsões dependem das informações mais recentes da série.

A escolha do valor mais apropriado para a constante de suavização pode ser determinada de forma interativa, ou seja, realizando uma varredura sobre um *grid* de valores dos parâmetros de suavização. O critério utilizado para a determinação dos valores ótimos para o vetor  $\alpha, \gamma$  será o de minimização da soma dos quadrados dos

erros de previsão, dado pela equação:  $\varepsilon = \sum_{t=1}^n (Y_t - \hat{Y}_t)^2$

### 2.3.2 – Alisamento Exponencial Holt-Winters

Este modelo é uma extensão do modelo de *Holt* para séries que apresentam tendência e sazonalidade. Existem duas formas para este modelo, a forma aditiva e a forma multiplicativa.

A forma aditiva do modelo, utilizada neste trabalho, baseia-se nas seguintes expressões para estimar, respectivamente, o nível, a tendência e a sazonalidade.

Nível .....  $n_t = \alpha(Y_t - f_{t-s}) + (1 - \alpha)(n_{t-1} + b_{t-1})$ , com  $0 \leq \alpha \leq 1$ ;

Tendência .....  $b_t = \beta(n_t - n_{t-s}) + (1 - \beta)b_{t-1}$  , com  $0 \leq \beta \leq 1$ ;

Sazonalidade .....  $f_t = \gamma(Y_t - n_t) + (1 - \gamma)f_{t-s}$  , com  $0 \leq \gamma \leq 1$ .

As previsões para instantes futuros ( $k$  instantes adiante) são obtidas através da seguinte equação:

$$\hat{Y}_{t+k} = k.b_t + f_{t+k-m.s}$$

Onde:  $m = 1$  para  $1 \leq k \leq s$

$m = 2$  para  $s \leq k \leq 2s$ .....etc.....

A forma multiplicativa do modelo assume uma interação entre a tendência e a sazonalidade. Dessa forma, estas duas componentes são agregadas na equação do modelo via o seu produto. Maiores detalhes podem ser obtidos em MORETTIN e TOLOI (2006).

### 2.3.3 - Comparação entre modelos de Alisamento Exponencial

De forma geral, são utilizadas três estatísticas de previsão para a comparação de modelos previsão, são elas:

MAPE – Erro percentual absoluto médio

$$MAPE = \left( \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \right) * 100\%$$

MAD - Erro absoluto médio ou Desvio absoluto médio

$$MAD = \left( \frac{1}{n} \sum_{t=1}^n |\hat{a}_t| \right)$$

MSD – Erro Quadrado médio ou Desvio quadrado médio

$$MSD = \left( \frac{1}{n} \sum_{t=1}^n \hat{a}_t^2 \right)$$

Onde “ $n$ ” significa quantos passos à frente se deseja prever.

Quanto menor o valor assumido por estas estatísticas, menor será o erro de previsão e, conseqüentemente, melhor será o ajuste do modelo.

## **2.4 - MÉTODOS AVANÇADOS DE PREVISÃO DE SÉRIES TEMPORAIS**

Conforme citado em MAKRIDAKIS (1985), no universo dos métodos de previsão de séries temporais encontram-se os modelos Autoregressivo e Médias Móveis (AR, MA e ARMA), modelos Autoregressivo Integrado de Médias Móveis (ARIMA), Filtros de Kalman e AEP, modelos ARARMA de Parzen, modelos ARMA Multivariáveis (MARMA), entre outros.

Os métodos que se classificam como avançados obtêm a previsão de algum valor futuro da série temporal pela combinação dos valores reais passados e/ou dos erros ocorridos.

Apenas os modelos AR, MA, ARMA e ARIMA serão apresentados, uma vez que foram amplamente utilizados para a modelagem das séries temporais deste trabalho.

### **2.4.1 Modelos Autoregressivo e de Médias Móveis**

MAKRIDAKIS (1985) descreve três procedimentos capazes de representar as observações de uma série temporal estacionária: modelo *Autoregressivo* (AR), modelo de *Médias Móveis* (MA) e o modelo *Autoregressivo e de Médias Móveis* (ARMA).

#### 2.4.1.1 - Modelo Autoregressivo

A especificação de um modelo Autoregressivo (AR) é dada pela equação:

$$Y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t \quad (20.1)$$

onde  $Y_t$  corresponde à observação da série temporal no tempo “ $t$ ”;

$\phi_p$  corresponde ao parâmetro do modelo AR de ordem “ $p$ ” e

$e_t$  representa o erro de eventos aleatórios que não podem ser explicados pelo modelo.

Caso as observações da série temporal possam ser representadas pela equação (20.1), será possível valores futuro da série MAKRIDAKIS (1985).

#### 2.4.1.2 - Modelo de Médias Móveis

Um modelo de Médias Móveis (MA) é definido conforme equação

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (20.2)$$

Onde:

$e_t$  representa o erro de eventos aleatórios que não podem ser explicados pelo modelo e

$\theta_q$  corresponde ao parâmetro do modelo MA de ordem  $q$ .

A interpretação da equação (20.2) é similar à equação (20.1), exceto pelo fato de que o valor previsto para a observação  $\alpha$  depende dos valores dos erros observados em cada período passado, ao invés das observações propriamente ditas.

### 2.4.1.3 - Modelo Autoregressivo e de Médias Móveis

MAKRIDAKIS (1985) especifica o modelo misto Autoregressivo e de Médias Móveis (ARMA) através da seguinte equação, como sendo a combinação dos modelos AR e MA.

$$Y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

Analisando esta equação é possível verificar que os modelos ARMA relacionam os valores futuros com as observações passadas, assim como também com os erros passados apurados entre os valores reais e os previstos.

### 2.4.2 Modelo de Box e Jenkins

O destaque atribuído ao modelo de George Box e Gwilyn Jenkins é um conseqüência de sua fundamentação teórica, aplicada a uma variedade de séries temporais (BOX e JENKINS, 1976).

O método de Box e Jenkins consiste na busca de um modelo ARIMA (Auto Regressive Integrate Moving Average) que represente o processo estocástico gerador da série temporal, a partir de um modelo ARMA aplicável na descrição de séries temporais estacionárias, estendendo esse conceito para séries temporais não-estacionárias.

Genericamente, um processo ARIMA  $(p,d,q)$  pode ser representado pela equação:

$$Y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$$

sendo  $Y_t = Y_t - Y_{t-d}$ ,

onde:

$\phi_p$  e  $\theta_q$  são os parâmetros dos processos Autoregressivo e de Média Móvel de ordem  $p$  e  $q$  (ARMA  $(p,q)$ );

$e_t$  corresponde ao erro de eventos aleatórios que não podem ser explicados pelo modelo e “ $d$ ” equivale ao grau de homogeneidade não-estacionária.

De acordo com BOX e JENKINS (1976), a estratégia para construção de um modelo ARIMA envolve uma abordagem iterativa que pode ser resumida da seguinte forma:

- (1) pressuposição de uma classe de modelos,
- (2) identificação do modelo,
- (3) estimação dos parâmetros do modelo,
- (4) verificação do modelo (ver se é adequado),
- (5) uso do modelo para previsão e controle.

Conforme MAKRIDAKIS (1985) o objetivo da *identificação* é determinar os valores de  $p$ ,  $d$  e  $q$  do modelo ARIMA  $(p,d,q)$ . Inicialmente, a série temporal  $Y_t$  é diferenciada para se obter uma série estacionária. Com isso, o processo fica reduzido a um modelo ARMA  $(p,q)$ . Em seguida, a ordem do processo ARMA é identificada pela análise dos coeficientes de autocorrelação e autocorrelação parcial. Ainda nesta etapa são efetuadas estimativas preliminares dos parâmetros do modelo identificado.

Após a identificação paramétrica do modelo, a *estimação* dos parâmetros desse modelo é efetuada. Os parâmetros do processo AR são estimados através de métodos de regressão. Caso o processo MA esteja envolvido, a estimação dos parâmetros é obtida pela aplicação de algoritmos de otimização não-linear.

#### 2.4.2.1 - Modelos não estacionários ARIMA $(p, d, q)$

Encontrar uma série estacionária não é tarefa fácil. Então, na maioria das séries será necessário aplicar o operador de diferença, que produz, a partir de um  $W_t$ , um

processo estacionário. Os modelos ARIMA são processos não estacionários cujas aplicações de diferenças sucessivas produzem um processo estacionário.

De acordo com MORETTIN e TOLOI (2004), se  $W_t = \Delta^d Y_t$  for estacionária, podemos representar  $W_t$  por um modelo ARMA ( $p, q$ ), ou seja:

$$\phi_q(B)W_t = \theta_p(B)a_t$$

Se  $W_t$  for uma diferença de  $Y_t$ , então  $Y_t$  é uma integral de  $W_t$ , daí tem-se que  $Y_t$  segue um modelo auto-regressivo, integrado, de médias móveis, ou Modelo ARIMA:

$$\phi_q(B)\Delta^d Y_t = \theta_p(B)a_t$$

de ordem ( $p, q, d$ ), onde  $p$  e  $q$  são as ordens de  $\phi(B)$  e  $\theta(B)$ , respectivamente.

Supõe-se que a  $d$ -ésima diferença da série  $Y_t$  pode ser representada por um modelo ARMA ( $p, q$ ) estacionário e inversível. Na maioria dos casos, utiliza-se  $d = 1$  ou  $d = 2$ , que correspondem a dois casos comuns de séries não-estacionárias homogêneas:

- i) Séries não-estacionárias quanto ao nível: oscilam ao redor de um nível médio durante algum tempo e depois saltam para outro nível temporário. Para torná-la estacionária, é suficiente tomar uma diferença.
- ii) Séries não-estacionárias quanto à inclinação: oscilam numa direção por algum tempo e depois mudam para outra direção temporária. Neste caso, para tornar a série estacionária, será necessário tomar duas diferenças.

A idéia de inversibilidade está relacionada com a possibilidade de escrevermos a série de ruído branco como saída de um filtro linear cuja entrada é a série estacionária  $\{W_t\}$ . Em linhas gerais, o processo  $\{W_t\}$  é inversível se:

$$a_t = \Psi(B)^{-1} w_t, \text{ fazendo } \Psi(B)^{-1} = \pi(B), \text{ temos: } a_t = \pi(B).w_t$$

$$\text{Onde: } \pi(B) = 1 - \pi_1 B - \pi_2 B^2;$$

Note que  $\pi(B) = \Psi^{-1}(B)$  e então os coeficientes  $\pi_1$ 's podem ser encontrados a partir dos  $\psi_1$ 's e vice versa. A condição de inversibilidade é aplicada aos coeficientes  $\pi_1$ , e é totalmente independente da condição de estacionariedade.

Dizemos que o processo  $\{W_t\}$  é inversível se  $\sum_{i=0}^{\infty} |\pi_i| < \infty$ , o que implica na convergência da série dentro do círculo unitário. Note que a condição de inversibilidade é análoga à condição de estacionariedade, mas a primeira é aplicada ao polinômio  $\pi(B)$ , enquanto a segunda refere-se ao polinômio  $\psi(B)$ .

#### 2.4.2.2 - Formas do modelo ARIMA

O modelo ARIMA pode ser escrito de três formas:

(i) Forma de Equação de Diferenças – Representada em termos dos valores prévios de  $Y_t$  e do valor atual. Sendo que este método é muito utilizado para fazer previsões.

$$Y_t = \xi_1 Y_{t-1} + \xi_2 Y_{t-2} \dots \xi_p Y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (24.1)$$

(ii) Formas de Choques Aleatórios – Representada em termos de valor atual e prévio de  $a_t$ . Sendo que este é um método conveniente para calcular a variância dos erros de previsão.

$$Y_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \psi(B)a_t \quad (24.2)$$

Desta equação obtém-se:

$$\xi(B)Y_t = \xi(B)\psi(B)a_t$$

Logo, os pesos  $\psi_j$  da Equação 24.1, estão associados aos parâmetros da Equação 24.2, ou seja,

$$(1 - \xi_1 B \dots \xi_p B^{p+d}) (1 + \psi_1 B + \psi_2 B^2 + \dots) = 1 - \theta_1 B \dots \theta_q B^q.$$

(iii) Forma Invertida

Representada em termos de valores prévios  $Y_t$  e do valor atual de  $a_t$ .

$$\text{obtem-se } \psi^{-1}(B)Y_i = a_i \quad (25.1)$$

$$\pi(B)Y_i = \left[ 1 - \sum_{j=1}^{\infty} \pi_j B^j \right] Y_i = a_i \quad (25.2)$$

$$\xi(B) = \theta(B) * \pi(B)$$

Portanto, os pesos  $\pi_j$  podem ser obtidos de (25.2), conhecendo-se os operadores de  $\xi(B)$  e acordo  $\theta(B)$ .

#### 2.4.2.3 - Identificação dos modelos ARIMA

O processo de identificação do Modelo ARIMA pode ser resumido em dois estágios:

1º Estágio: escolha do grau de diferenciação “d”, isto é transformar  $Y_t$  em  $W_t$ .

2ª Estágio: escolha do grau dos polinômios “p”, “q” do modelo ARMA a ser aplicado à série  $W_t$ .

A identificação da ordem do modelo mencionado, independente do estágio, se faz através das funções de autocorrelação (FAC) e autocorrelação parcial (FACP).

- Grau de diferença, escolhendo o “d”:

A identificação do grau de diferenças “d” do Operador  $\nabla^d$  aplicado  $Y_t$ , para modelos não estacionários, é feita através da análise da função de autocorrelação estimada ( $\hat{\rho}_k$ ) e a função de autocorrelação teórica ( $\rho_k$ ).

Se a Função de Autocorrelação mostra decaimento muito lento, o modelo não é estacionário, sendo necessária sua diferenciação a fim de torná-lo estacionário.

- Grau dos polinômios, escolhendo “ $p$ ” e “ $q$ ”:

A identificação de “ $p$ ” e “ $q$ ” é feita comparando o comportamento dos estimadores das autocorrelações  $(\rho_k)$  e das autocorrelações parciais  $\left(\phi_{kk}\right)$  com as correspondentes funções teóricas.

Neste trabalho utilizamos as Funções de Autocorrelação e Autocorrelação Parcial para identificação do modelo. Porém, existem alguns outros métodos para identificar os modelos ARIMA. Por exemplo: Função de Autocorrelação Inversa, Arrays R&S ou Transformações G e Função de Autocorrelação Estendida. Maiores detalhes podem ser encontrados em MORETTIN e TOLOI (2004).

#### 2.4.2.4 - Estimação de modelos ARIMA

Considerando um modelo ARIMA  $(p, q, d)$ , inserimos os seus parâmetros  $p+q+1$  no vetor  $\zeta = (\phi, \theta, \sigma_a^2)$ , onde:  $\phi = (\phi_1, \dots, \phi_p)$ ,  $\theta = (\theta_1, \dots, \theta_q)$ , supondo que  $\mu_w = 0$  quando  $d > 0$ . Caso contrário,  $\mu$  é incluído como mais um parâmetro a ser estimado e teremos  $p+q+2$  parâmetros.

Para estimar o vetor de parâmetros  $\zeta$ , o método empregado será o de máxima verossimilhança:

Dada as “ $n$ ” observações de  $Y_1, \dots, Y_n$  consideramos a função de verossimilhança  $L(\zeta | Y_1, \dots, Y_n)$  como função de  $\zeta$ . Os estimadores de máxima verossimilhança (EMV) de  $\zeta$  serão os valores que maximizam a função de verossimilhança,  $L$ , ou de forma equivalente, a função log verossimilhança  $\ell = \log L$ .

Para determinarmos os EMV trabalharemos com a suposição de que o processo  $a_t$  está “normalmente distribuído”, ou seja, para cada  $t, a_t \sim N(0, \sigma_a^2)$ . Nestas condições, os EMV serão aproximadamente estimadores de mínimos quadrados (EQM).

#### 2.4.2.5 - Validação dos modelos ARIMA

##### a) análise de resíduos:

Após construir a série de resíduos estimados  $\hat{a}_t$ , através do modelo escolhido, verificamos se os mesmos atendem as suposições de normalidade, média zero, variância constante e independência. Este conjunto particular de suposições é chamado de processo ruído branco.

Caso a série  $\hat{a}_t$  seja estacionária, calculamos as Funções de Autocorrelação e Autocorrelação Parcial. Se a análise das FAC e FACP indicarem um processo ruído branco, o modelo escolhido para  $Y_t$  poderá ser utilizado para fins de previsão ou controle. Caso contrário, a análise de resíduos poderá ser utilizada para identificar outro modelo mais adequado a série.

##### b) previsão

Para que possamos proceder às previsões, assumiremos que  $W_t = (1 - B)^d Y_t$  é estacionário e inversível e os parâmetros do modelo são conhecidos. A previsão se torna o cálculo do valor esperado de uma futura observação, condicionada aos valores passados e ao valor presente da variável.

$$\hat{Y}_t(r) = E(Y_{t+r} | Y_t, Y_{t-1}, \dots)$$

E o valor de  $Y$  no tempo  $t + r$  é obtido pela equação do modelo ARIMA.

Conforme Souza e Camargo (2004), podemos expressar a previsão de, pelo menos, duas formas: através da equação de diferenças, na forma de choques aleatórios.

c) estatísticas para comparação dos modelos: erro de previsão

O Erro de Previsão é uma medida também utilizada para escolher o modelo mais apropriado para análise do conjunto de dados. Este critério sugere que o melhor modelo é o que apresenta menor erro de previsão.

A fórmula matemática para do erro de previsão é:

$$MSE_p = \frac{\sum_{t=1}^k (Y_t - \hat{Y}_t)^2}{k}$$

onde k é o número de previsões realizadas.

### **3 ESCOLHA E AJUSTE DE MODELOS**

Neste t3pico ser3 apresentada cada s3rie temporal separadamente e feitos a escolha e o ajuste do modelo que melhor reproduz a s3rie.

#### **3.1 - DADOS (an3lise preliminar)**

Para a realiza33o do presente trabalho, utilizamos diversas s3ries temporais relativas aos impostos arrecadados pelo munic3pio de Belo Horizonte, compreendendo o per3odo de janeiro de 2002 a abril de 2009.

A s3rie temporal denominada IMP PREDIAL 3 formada pelos recolhimentos efetuados pelos contribuintes do Imposto Predial relativo aos im3veis edificados em Belo Horizonte, levando-se em conta o m3s em que o recolhimento do imposto foi efetuado.

Por sua vez, a s3rie temporal denominada IMP TERRIT. URBANO 3 formada pelos recolhimentos efetuados pelos contribuintes do Territorial Urbano relativo aos im3veis n3o edificados (lotes vagos) em Belo Horizonte por m3s em que o recolhimento do imposto foi efetuado.

As Figuras 1 e 2, respectivamente, apresentam o comportamento de duas s3ries temporais relativas ao Imposto Predial (IMP. PREDIAL) e ao Imposto Territorial Urbano (IMP. TERRIT. URBANO). Em ambos os casos, verifica-se que h3 sazonalidade. Esta sazonalidade pode ser explicada pelo incentivo que 3 dado pelo governo municipal para o recolhimento antecipado de todas as parcelas no m3s de janeiro de cada ano. O imposto 3 cobrado em 12 parcelas que vencem mensalmente. O governo do munic3pio incentiva o recolhimento antecipado, concedendo descontos sobre o valor total do tributo, criando um efeito sazonal no m3s de janeiro de cada ano.

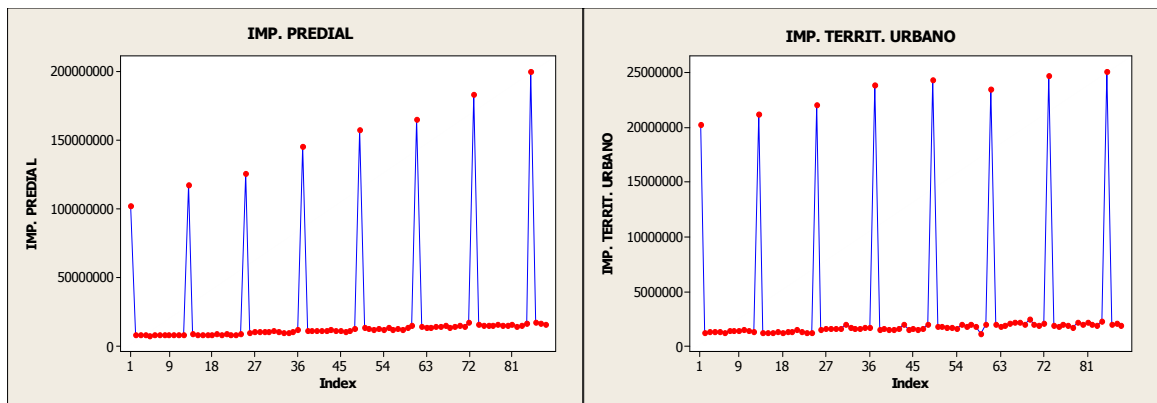


Figura 1: Gráfico da série Imposto Predial

Figura 2: Gráfico da série Imposto Territorial Urbano

A série temporal denominada ITBI (Figura 3) é formada pelos recolhimentos efetuados pelos contribuintes do Imposto sobre Transmissão de Bens Imóveis. Este imposto incide sobre as transações comerciais imobiliárias realizadas em Belo Horizonte relativas a imóveis, edificados ou não, situados no território do município, por mês em que o recolhimento do imposto foi efetuado.

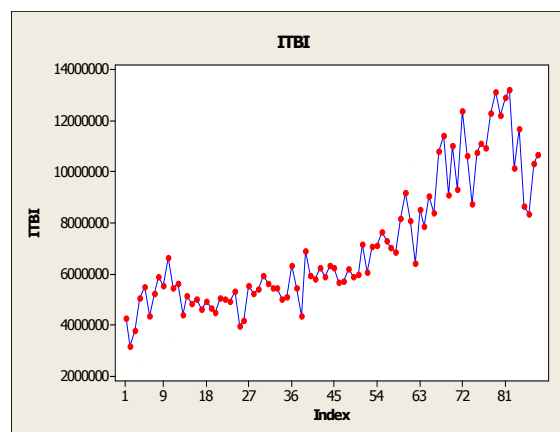


Figura 3: Gráfico da série Imposto sobre Transmissão de Bens Imóveis

Os recolhimentos relativos ao ITBI não respeitam um vencimento mensal específico, ou seja, não segue o padrão de vencimento do Imposto Predial ou Territorial. Os valores do imposto são recolhidos à medida que as operações de compra e venda ocorrem, significando que o crescimento demonstrado na série se refere ao incremento das transações comerciais ou ao aumento médio dos preços dos imóveis transacionados.

O Imposto Sobre Serviços de Qualquer Natureza (ISSQN) é devido nas operações de prestação de serviço, conforme definido em lei específica, e recolhido no mês subsequente ao evento. Sobre o valor do serviço prestado são aplicadas alíquotas que variam conforme a natureza do serviço, entre 2% e 5%, o valor resultante é o imposto devido.

O legislador achou por bem criar duas modalidades diferentes de recolhimento para o ISS. Numa delas, o próprio prestador do serviço é quem faz o recolhimento do imposto devido sobre a operação, na outra, o tomador do serviço é quem estará responsabilizado por descontar do valor pago ao prestador a parte relativa ao imposto e, ainda, recolher ao cofre do município. Esta prática é denominada “retenção na fonte”.

As Figuras 4 e 5 apresentam o comportamento das séries denominadas ISS RETIDO e ISS PRÓPRIO. A primeira se refere ao ISS recolhido pelos tomadores de serviço e a outra o ISS recolhido pelo prestador de serviços. A clara tendência de crescimento demonstrada nas séries se refere ao incremento das transações comerciais abrangidas pela incidência do imposto.

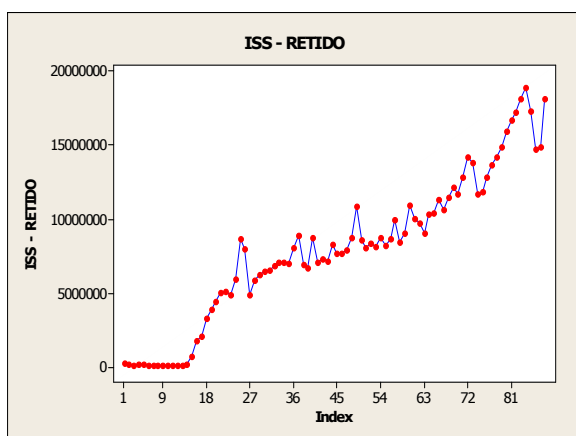


Figura 4: Gráfico da série ISS Retido na Fonte

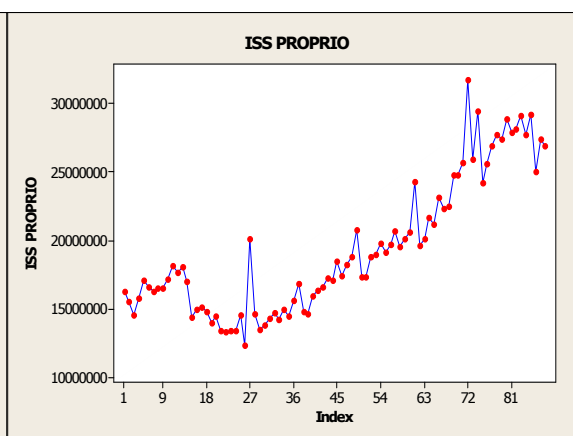


Figura 5: Gráfico da série ISS Próprio

As duas séries do ISS próprio e retido foram apresentadas separadas pois são elementos distintos da escrituração contábil de receitas municipais, mas na prática, tratam-se de séries complementares.

As operações sujeitas ao ISSQN retido ou próprio são as mesmas, o que as diferencia é a forma de recolhimento. Conforme explicado nos parágrafos anteriores. Para o objeto deste estudo o que importa é a identificação de um modelo de previsão. Neste sentido, as séries descritas separadamente foram somadas e estudadas sobre a mesma ótica.

Entendemos que este exercício não afetará os resultados, pelo contrário, facilitará o entendimento da dinâmica de operações comerciais sujeitas ao imposto, sem se preocupar com o detalhe de quem é o responsável pelo recolhimento, que é irrelevante neste estudo.

Do mesmo modo que o ISS PRÓPRIO e o ISS-RETIDO, o ISS - AUT. LIBERAIS é uma outra modalidade de recolhimento do Imposto sobre Serviços. A série temporal denominada ISS - AUT. LIBERAIS é formada pelos recolhimentos efetuados pelos contribuintes do Imposto sobre Serviços cadastrados como prestadores de serviço autônomos ou como profissionais liberais.

Nesta modalidade o imposto é apurado em valores fixos, não se aplica uma alíquota percentual sobre o valor do serviço prestado. O contribuinte paga um valor fixo mensal conforme definido em lei. O vencimento do imposto pode ser trimestral ou mensal dependendo da condição em que se deu o cadastramento do prestador de serviços.

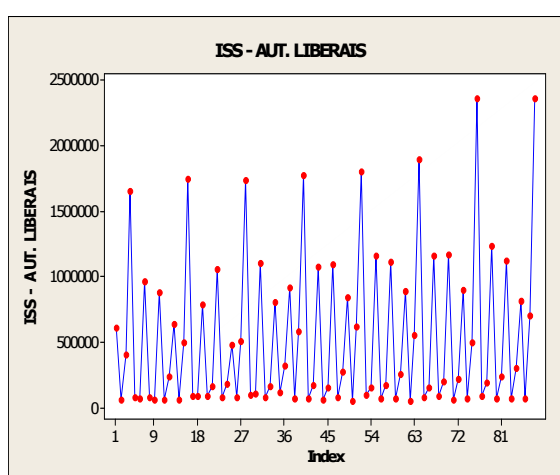


Figura 6: Gráfico da série ISS – Autônomos e Liberais

O gráfico da Figura 6 apresenta o comportamento da série ao longo do tempo. Pela análise visual do gráfico, parece haver sazonalidade que deve se referir ao vencimento trimestral do imposto. Há uma concentração de recolhimentos a cada três meses. O crescimento demonstrado na série se refere ao incremento do número de profissionais inscritos no cadastro de prestadores do município e não necessariamente ao aumento das transações comerciais abrangidas pela incidência do imposto.

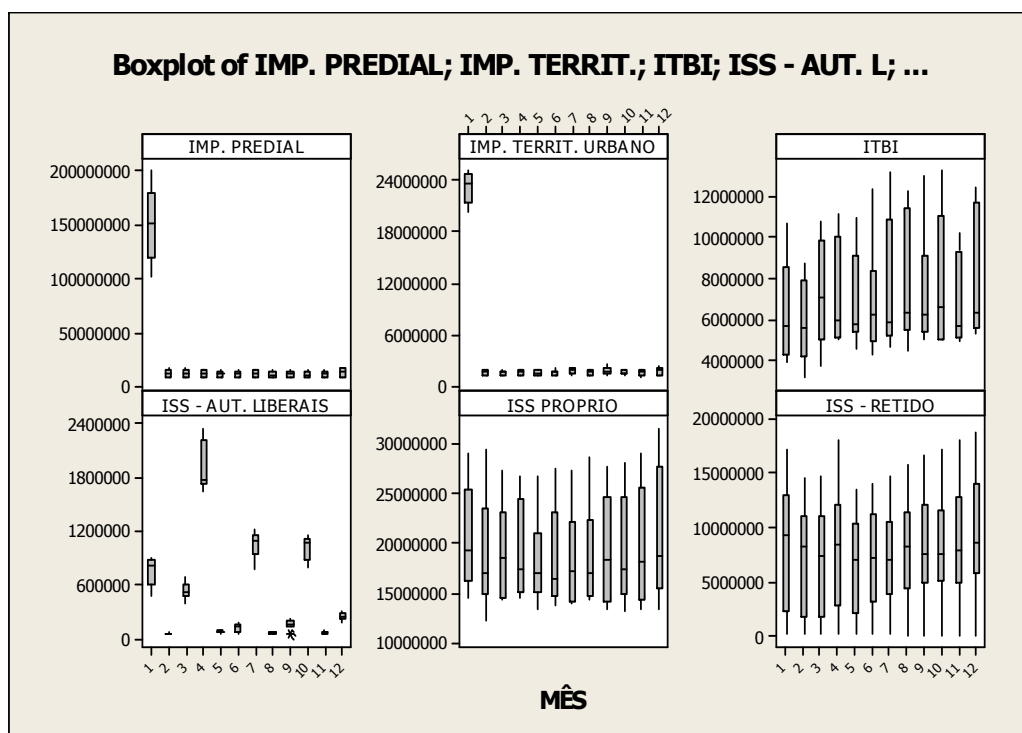


Figura 7: Box-Plot mensal de todas as séries estudadas

Ao analisarmos a Figura 7, referente ao Box-Plot mensal de todas as séries estudadas neste trabalho, verificamos a existência de forte sazonalidade em algumas das séries, principalmente, aquelas que tratam dos impostos predial e territorial urbanos.

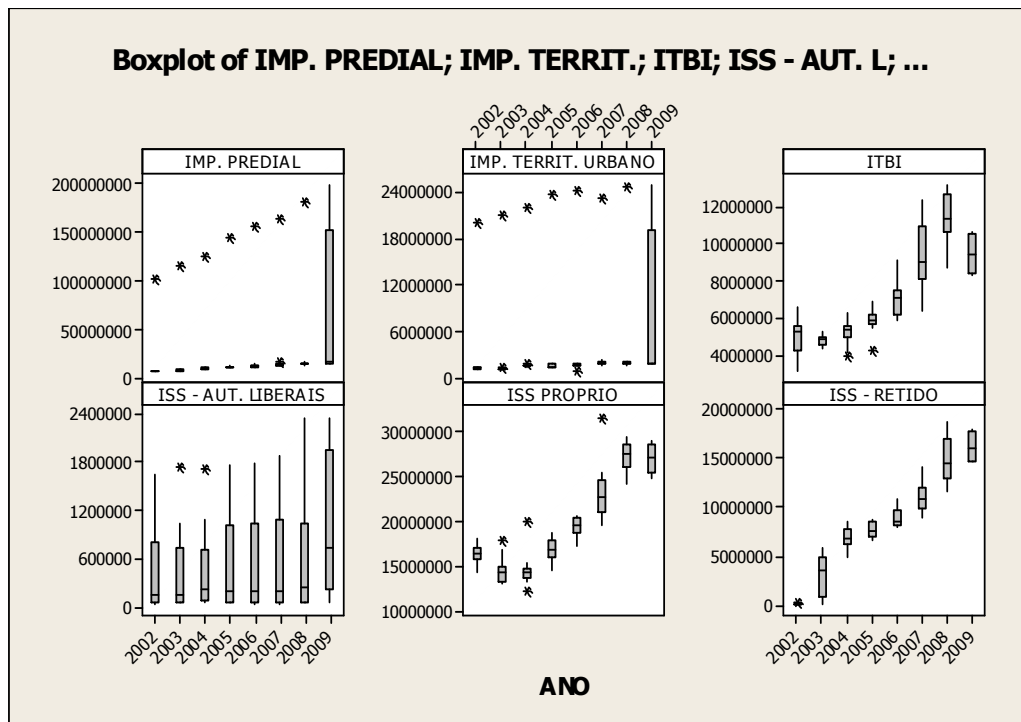


Figura 8: Box-Plot anual de todas as séries estudadas

A Figura 8, referente ao Box-Plot anual de todas as séries estudadas neste trabalho, nos mostra que as séries estudadas não são estacionárias, já que as medianas não se encontram em nenhum momento alinhadas. Verificamos a presença de out-liers que deverão ser objeto de investigação mais profunda no decorrer da apresentação do trabalho.

Quadro 1: Estatísticas descritivas das séries temporais analisadas

| Variable            | Mean     | StDev     | CoefVar | Minimum  | Q1       | Median   |
|---------------------|----------|-----------|---------|----------|----------|----------|
| IMP. PREDIAL        | 24374234 | 41031863  | 168,34  | 7884933  | 9339302  | 12173303 |
| IMP. TERRIT. URBANO | 3631351  | 6214516   | 171,14  | 1101731  | 1481050  | 1719294  |
| ITBI                | 7120444  | 2565610   | 36,03   | 3162940  | 5232902  | 6141079  |
| ISS - AUT. LIBERAIS | 526350   | 580854    | 110,36  | 46718    | 78858    | 226700   |
| ISS PROPRIO         | 19375959 | 4996968   | 25,79   | 12355398 | 14969737 | 17849402 |
| ISS - RETIDO        | 7966041  | 5027798   | 63,12   | 103671   | 4920603  | 8105113  |
| Variable            | Q3       | Maximum   |         |          |          |          |
| IMP. PREDIAL        | 15163409 | 199960097 |         |          |          |          |
| IMP. TERRIT. URBANO | 1994168  | 25053403  |         |          |          |          |
| ITBI                | 8728793  | 13231168  |         |          |          |          |
| ISS - AUT. LIBERAIS | 888345   | 2358174   |         |          |          |          |
| ISS PROPRIO         | 22970554 | 31727944  |         |          |          |          |
| ISS - RETIDO        | 11217307 | 18872792  |         |          |          |          |

Obs.: cálculos realizados utilizando recursos do programa Minitab 15

O Quadro 1 apresenta as estatísticas descritivas das séries estudadas neste trabalho. O coeficiente de variação demonstra que os dados se dispersam muito em relação à média para todas as séries. O destaque é para a série IMP. TERRIT. URBANO com o maior índice de variação. O menor índice é verificado na série ISS PROPRIO.

Adiante apresentaremos as séries originais convertidas em Logaritmo base 10 (Log10). Esta alternativa foi utilizada para reduzir a variabilidade dos dados e facilitar a análise visual dos gráficos e informações utilizados neste trabalho.

### 3.2 - ANÁLISE DESCRITIVA DOS DADOS - IMPOSTO PREDIAL

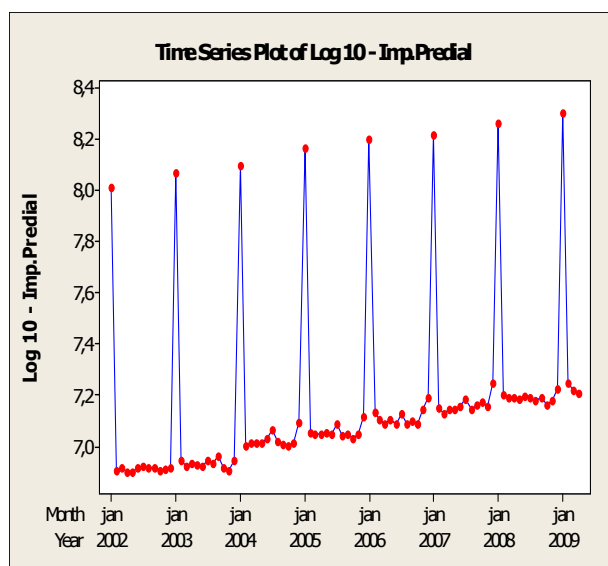


Figura 9: Gráfico da série Log10 Imposto Predial

A Figura 9 apresenta o gráfico da série do Imposto Predial (Log10 Imp.Predial) para o período de janeiro de 2002 a abril de 2009. Observamos que a série apresenta crescimento ao longo de todo o período analisado e que nos meses de janeiro há um aumento considerável do valor recolhido do imposto, evidenciando sazonalidade.

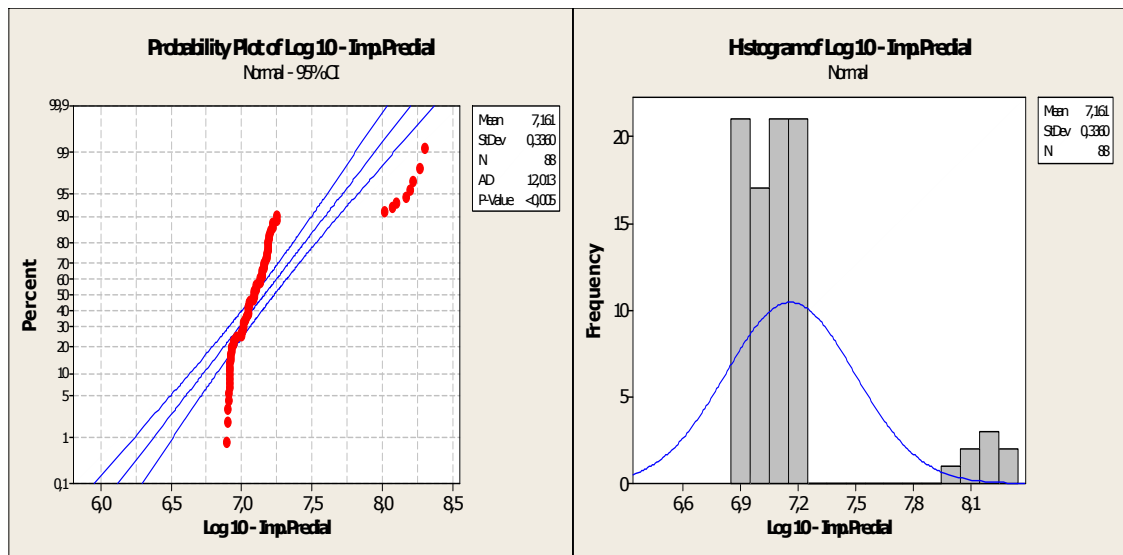


Figura 10: Gráfico de probabilidade e histograma da série Log10 Imposto Predial

A análise dos gráficos da Figura 10 evidencia que os dados da série Log10 Imp.Predial não se aproximam de uma Distribuição Normal, já que o p-valor apresentado é menor que 0,05.

### 3.2.1 - Ajuste do modelo de regressão

Para o caso das receitas municipais com o recolhimento do Imposto Predial, optamos por ajustar um modelo de regressão. Este modelo tem como variável resposta o Log10 do Imposto Predial e como variáveis preditoras um contador que identifica a sequência da série temporal (tempo), o fato de ser a predição relativa ao mês de janeiro ou não (variável jan) e a variável “jan” multiplicada pelo tempo (janXtempo).

**Equação de regressão:**

$$\text{Log10 ImpPredial} = 6,88 + 0,00406(\text{tempo}) + 1,14(\text{jan}) - 0,0007(\text{janXtempo})$$

Tabela 1 – Estimativas dos parâmetros da regressão para a série Log10 Imposto Predial

| Variável  | Coefficiente | Desvio padrão | Teste T | P-valor |
|-----------|--------------|---------------|---------|---------|
| Constante | 6,87978      | 0,00577       | 1192,32 | 0,000   |
| Tempo     | 0,0040609    | 0,0001126     | 36,08   | 0,000   |
| Jan       | 1,14091      | 0,01761       | 64,79   | 0,000   |
| JanXtempo | -0,0007001   | 0,0003448     | -2,03   | 0,046   |

Considerando os dados da Tabela 1, constatamos que os parâmetros são significativos para o modelo, ao nível de significância de 5%.

O modelo ajustado pela regressão linear simples apresenta um  $R^2$  de 99,5%;  $R^2$  ajustado de 99,4%; Valor do teste F de 5.068,16; p-valor igual a 0,000 e estatística de Durbin-Watson 1,33344.

Para atestarmos a suposição de independência dos resíduos foi aplicado o teste de Durbin-Watson, que é utilizado para detectar autocorrelações de primeira ordem nos resíduos, sendo que:

$H_0$  :  $\rho = 0$  - existe independencia;

$H_1$  :  $\rho \neq 0$  - existe dependencia.

Utilizando a tabela de valores críticos da estatística de Durbin-Watson para  $\alpha=0,05$  e  $n= 88$  e  $k=3$ , temos:  $d_L = 1,57$  e  $d_U = 1,72$ .

Uma vez que  $d=1,33344 < 1,57$  ( $d < d_L$ ) é rejeitada a hipótese de independência dos resíduos, pois há evidências que os erros estão positivamente correlacionados.

Ainda testando a validade do modelo, analisamos os resíduos apresentados na regressão, que são mostrados na Figura 11.

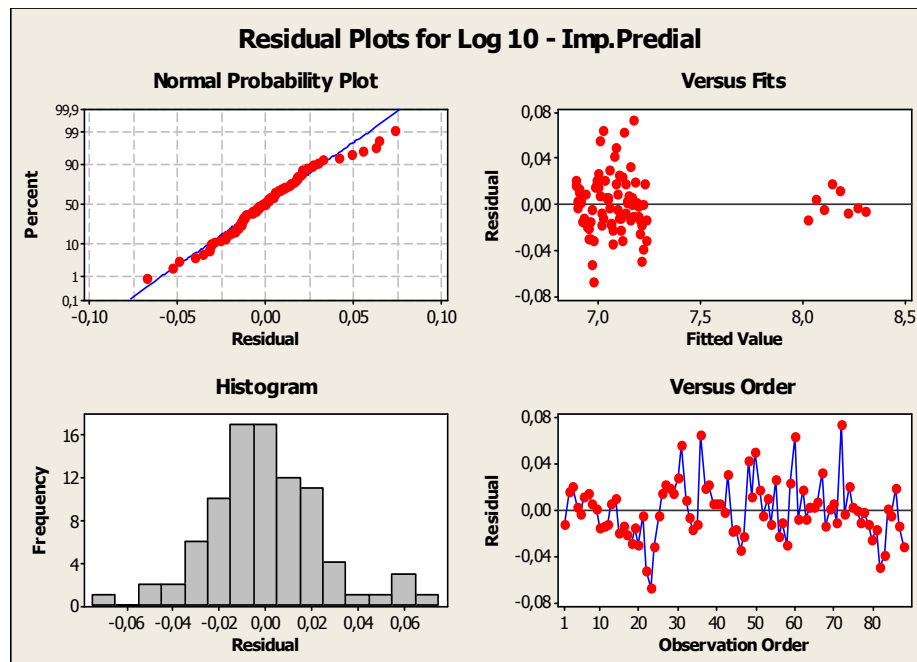


Figura 11: Resíduos do ajuste de Regressão para a série Log10 do Imposto Predial

Na validação do modelo pelos resíduos apresentados na Figura 11, podemos inferir que se tratam de resíduos homocedásticos e estão distribuídos normalmente, conforme teste de normalidade apresentado na Figura 12. Contudo, estão correlacionados, conforme estatística de Durbin-Watson.

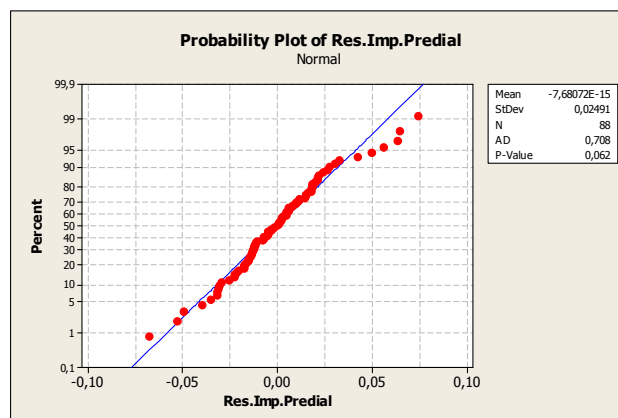


Figura 12: Teste de Normalidade para os Resíduos do ajuste de Regressão para o Log10 do Imposto Predial

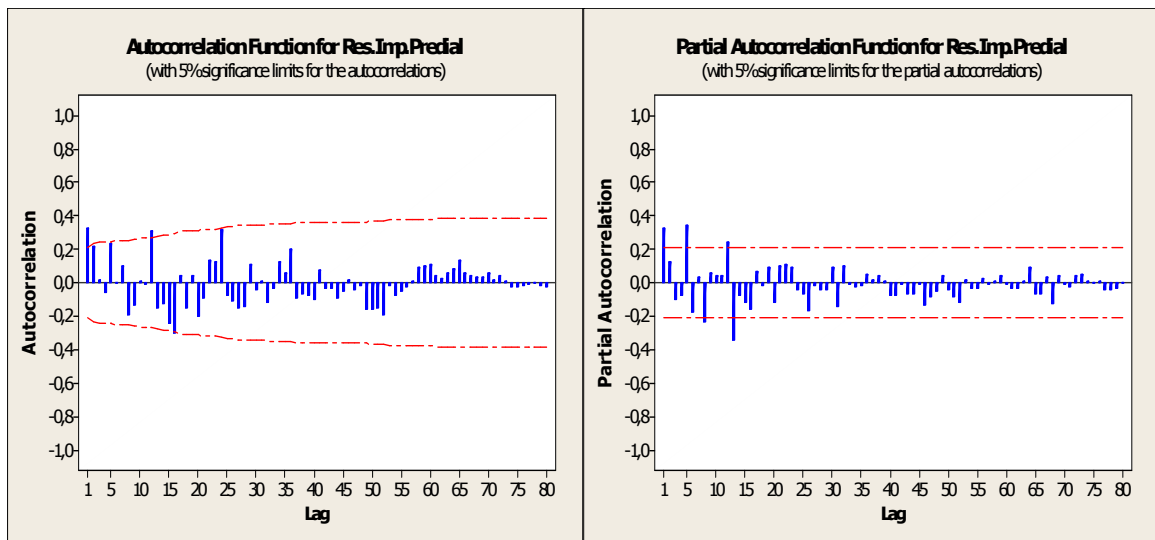


Figura 13: FAC e FACP para os Resíduos do ajuste de Regressão para a série Log10 do Imposto Predial

A Figura 13 apresenta a Função de Correlação e Função de Correlação Parcial dos resíduos do Imposto Predial e nos revela tratar-se de um processo de médias móveis com sazonalidade, demonstrando a dependência temporal entre os resíduos.

Considerando que o pressuposto de independência dos resíduos foi violado, precisamos buscar alternativas para superar este problema. Assim, inserimos mais uma variável preditora no modelo de regressão. Essa variável que representa os resíduos oriundos da estimativa inicial é dada pela seguinte fórmula:  $e_t = y_t - \hat{y}_t$ .

O teste de Durbin Watson se restringe a correlações de primeira ordem e que a FAC e FACP (vide Figura 13) não nos dá uma visão clara do processo, se se trata de um processo auto-regressivo ou de médias móveis, adotamos a defasagem de um passo atrás. Esta proposta evidencia que o valor observado em  $t_{(0)}$  guarda correlação com o valor observado no tempo  $t_{(-1)}$ . Ou seja o erro observando em  $t_{(-1)}$  será utilizado como variável preditora para o valor estimado em  $t_{(0)}$ .

### 3.2.2 – Ajuste do modelo de regressão com inserção do erro de previsão

Este modelo tem como variável resposta o Log10 do Imposto Predial e como variáveis predictoras um contador que identifica a sequencia da série temporal

(tempo), o fato de ser a predição relativa ao mês de janeiro ou não (variável jan), a variável “jan” multiplicada pelo “tempo” (janXtempo) e a variável resíduo<sub>(t-1)</sub>.

### Equação de regressão:

$$\text{Log 10 - Imp.Predial} = 6,88 + 0,00408 \text{ tempo} + 1,15 (\text{jan}) - 0,00117 (\text{janXtempo}) + 0,392 \text{ resíduo (t-1)}$$

Tabela 2 – Estimativas dos parâmetros da regressão para a série Log10 Imposto Predial

| Variável                 | Coefficiente | Desvio padrão | Teste T  | P-valor |
|--------------------------|--------------|---------------|----------|---------|
| Constante                | 6,87977      | 0,00544       | 1.264,94 | 0,000   |
| Tempo                    | 0,0040797    | 0,0001020     | 38,40    | 0,000   |
| Jan                      | 1,15359      | 0,02126       | 54,27    | 0,000   |
| JanXtempo                | -0,0011672   | 0,0004010     | -2,91    | 0,005   |
| Resíduo <sub>(t-1)</sub> | 0,3922       | 0,1132        | 3,47     | 0,001   |

Considerando os dados da Tabela 2, constatamos que os parâmetros são significativos para o modelo, ao nível de significância de 5%.

O modelo ajustado pela regressão linear simples apresenta um R<sup>2</sup> de 99,5%; R<sup>2</sup> ajustado de 99,5%; Valor do teste F de 3.961,67; p-valor igual a 0,000 e estatística de Durbin-Watson (d) 1,9564.

Utilizando a tabela de valores críticos da estatística de Durbin-Watson para  $\alpha=0,05$  e  $n= 88$  e  $k=4$ , temos:  $d_L = 1,57$  e  $d_U = 1,75$ .

Uma vez que  $(4 - d) = 2,0436 > d_U = 1,75$  e que  $d = 1,9564 > d_U = 1,75$ , não há evidências da existência de correlações positivas ou negativas entre os erros, assim, não é rejeitada a hipótese de independência. Podemos, pois, admitir que os erros são independentes.

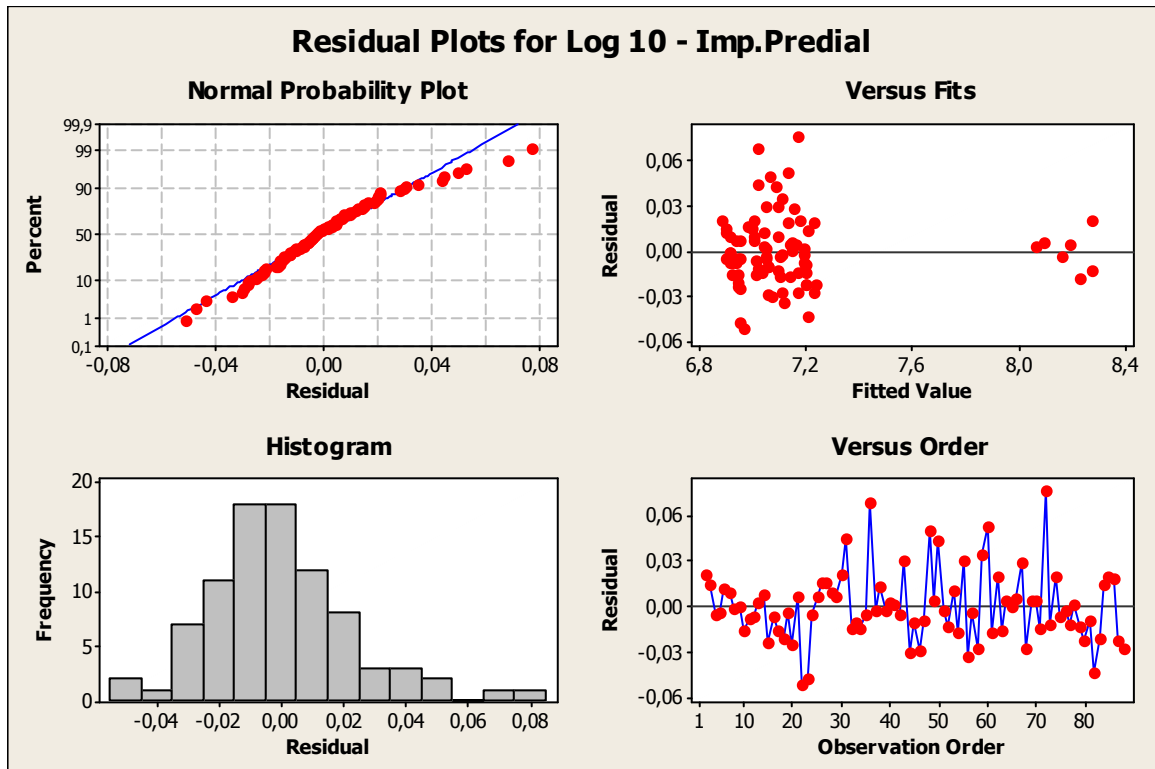


Figura 14: Resíduos do ajuste de Regressão para a série Log10 do Imposto Predial

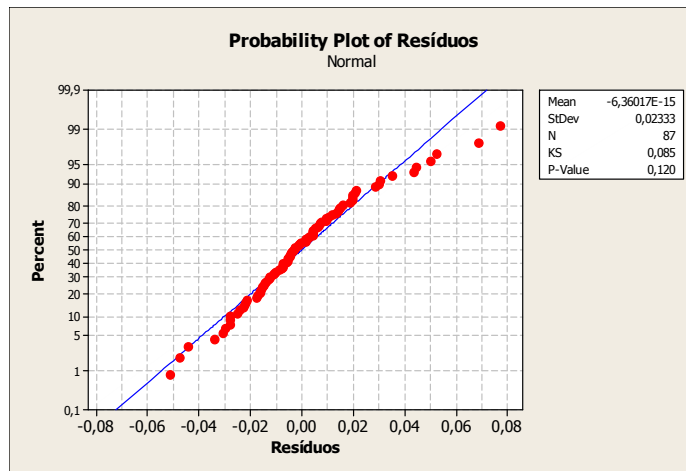


Figura 15: Teste de normalidade para os resíduos do ajuste de Regressão da série Log10 do Imposto Predial

Na validação do modelo pelos resíduos apresentados nas Figura 14, podemos inferir que se tratam de resíduos homocedásticos, distribuídos normalmente, conforme teste de normalidade apresentado na Figura 15 e não correlacionados, conforme estatística de Durbin-Watson.

### 3.2.3 - Previsão usando o modelo de regressão ajustado

Tabela 3: Previsões (1 passo a frente) maio a setembro de 2009 – Imposto Predial, valores em Reais (R\$)

| <b>Mês</b> | <b>Previsão limite inferior</b> | <b>Previsão</b> | <b>Previsão limite superior</b> | <b>Valor observado</b> | <b>Erro de previsão</b> |
|------------|---------------------------------|-----------------|---------------------------------|------------------------|-------------------------|
| maio       | 15.189.805,46                   | 17.009.983,10   | 19.042.667,67                   | 16.298.757,61          | -711.225,49             |
| junho      | 15.792.676,12                   | 17.668.512,98   | 19.767.159,71                   | 16.090.016,08          | -1.578.496,90           |
| julho      | 15.940.197,57                   | 17.835.282,80   | 19.955.669,38                   | 17.413.473,92          | -421.808,88             |
| agosto     | 16.089.064,54                   | 18.003.626,73   | 20.146.017,47                   | 16.162.680,53          | -1.840.946,20           |
| setembro   | 16.239.289,09                   | 18.173.559,62   | 20.338.222,16                   | 16.406.477,86          | -1.767.081,76           |

Para fins de cálculo dos valores projetados, utilizamos o valor do erro registrado em abril/2009 para a estimativa do mês de maio/2009. Para os demais meses não foram utilizados valores de erro como variável preditora, posto não haver valores específicos e a média do erro ser igual a zero.

A Tabela 3 apresenta as estimativas um passo a frente para o período de Maio a Setembro de 2009 relativas ao Log 10 do Imposto Predial. Os valores em logaritmo foram ajustados para a moeda corrente. Em termos absolutos, o maior erro da estimativa se deu em agosto/2009 e o menor em julho/2009. Em termos percentuais, os erros representam respectivamente 11,39% e 2,42%.

As previsões dos limites inferior e superior foram construídas observando-se um intervalo de confiança de 95%, conforme metodologia explicitada em MONTGOMERY e RUNGER (2003, p.246).

### 3.3 - ANÁLISE DESCRITIVA DOS DADOS - IMPOSTO TERRITORIAL URBANO

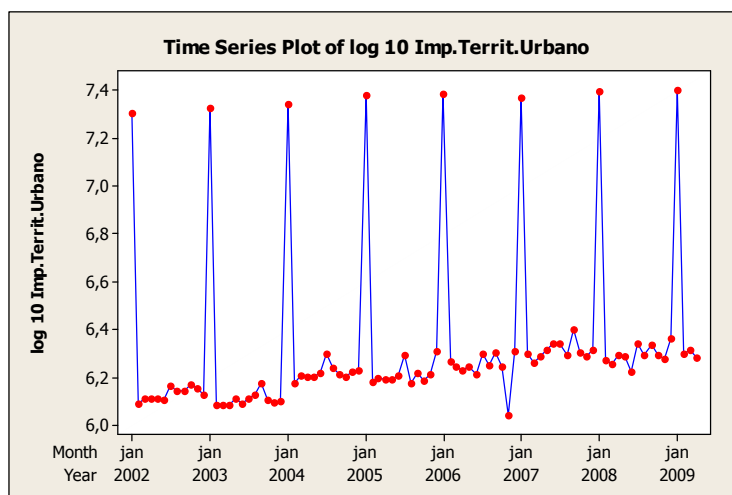


Figura 16: Gráfico da série Log 10 do Imposto Territorial Urbano

A Figura 16 apresenta o gráfico da série Log10 Imposto Territorial Urbano (Imp.Territ.Urbano) para o período de janeiro de 2002 a abril de 2009. Observamos que a série apresenta crescimento ao longo de todo o período analisado e que nos meses de janeiro há um aumento considerável do valor recolhido do imposto, evidenciando sazonalidade.

O gráfico da Figura 16 também nos permite evidenciar um dado que discrepante no mês novembro de 2006. Vejamos que enquanto a tendência é de crescimento no período, o valor de recolhimento do imposto, decresce no citado mês, inexistindo justificativas conhecidas para este valor.

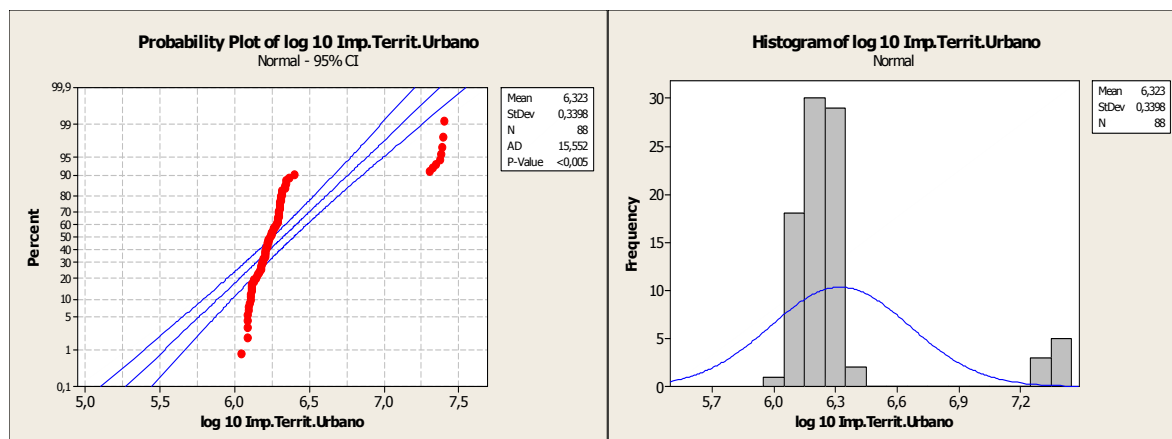


Figura 17: Gráfico de probabilidade e histograma da série Log 10 do Imposto Territorial Urbano

A análise dos gráficos da Figura 17 evidencia que os dados da série logarítmica do Imposto Territorial Urbano não se aproximam de uma Distribuição Normal, já que o p-valor apresentado é menor que 0,05.

### 3.3.1 - Ajuste do modelo de regressão

Para o caso das receitas municipais com o recolhimento do Imposto Territorial Urbano, optamos por ajustar um modelo de regressão. Este modelo tem como variável resposta o Imposto Territorial Urbano e como variáveis predictoras um contador que identifica a sequencia da série temporal (tempo), o fato de ser a predição relativa ao mês de janeiro ou não (variável jan), a variável dummy “janeiro” multiplicada pelo tempo (janXtempo) e a variável novembro de 2006 (Nov\_2006). Esta última variável foi introduzida para dar tratamento adequado ao dado discrepante encontrado no mês de novembro de 2006.

#### Equação de regressão:

$$\text{Log10 Imp.Territ.Urbano} = 6,10 - 0,219(\text{Nov\_2006}) + 0,00273(\text{tempo}) + 1,22(\text{jan}) - 0,00165(\text{janXtempo})$$

Tabela 4 – Estimativas dos parâmetros da regressão para a série Log10 Imposto Territorial

| Variável  | Coefficiente | Desvio padrão | Teste T | P-valor |
|-----------|--------------|---------------|---------|---------|
| Constante | 6,100019     | 0,00897       | 680,01  | 0,000   |
| Nov_2006  | -0,21907     | 0,03974       | -5,51   | 0,000   |
| Tempo     | 0,0027282    | 0,0001754     | 15,56   | 0,000   |
| Jan       | 1,21554      | 0,02738       | 44,40   | 0,000   |
| JanXtempo | -0,0016481   | 0,0005363     | -3,07   | 0,003   |

Considerando os dados da Tabela 4, constatamos que os parâmetros são significativos para o modelo, ao nível de significância de 5%.

O modelo ajustado pela regressão linear simples apresenta um  $R^2$  de 98,7%;  $R^2$  ajustado de 98,7%; Valor do teste F de 1.595,83; p-valor igual a 0,000 e estatística de Durbin-Watson 1,25374.

Utilizando a tabela de valores críticos da estatística de Durbin-Watson para  $\alpha=0,05$  e  $n= 88$  e  $k=3$ , temos:  $d_L = 1,57$  e  $d_U = 1,72$ .

Uma vez que  $d=1,25374 < 1,57$  ( $d < d_L$ ) é rejeitada a hipótese de independência dos resíduos, pois há evidências que os erros estão positivamente correlacionados.

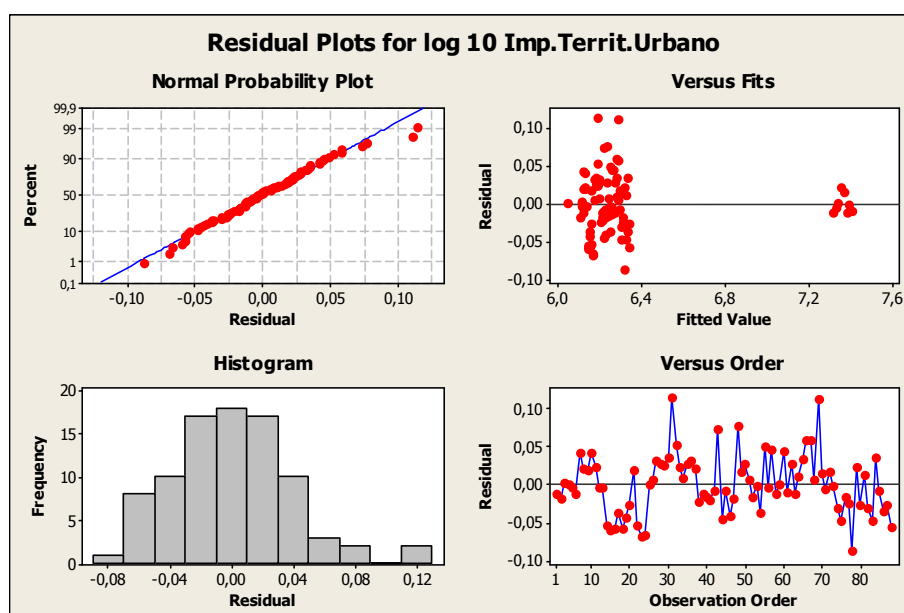


Figura 18: Resíduos do ajuste de Regressão para série Log 10 do Imposto Territorial Urbano

Na validação do modelo pelos resíduos apresentados na Figura 18, podemos inferir que se tratam de resíduos homocedásticos e estão distribuídos normalmente, conforme teste de normalidade apresentado na Figura 19. Contudo, estão correlacionados, conforme estatística de Durbin Watson.

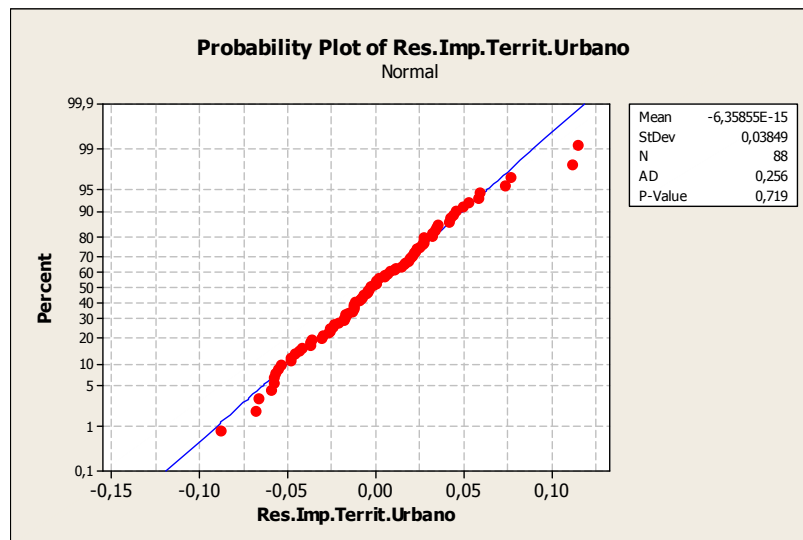


Figura 19: Teste de Normalidade para os Resíduos do ajuste de Regressão para série Log 10 do Imposto Territorial Urbano

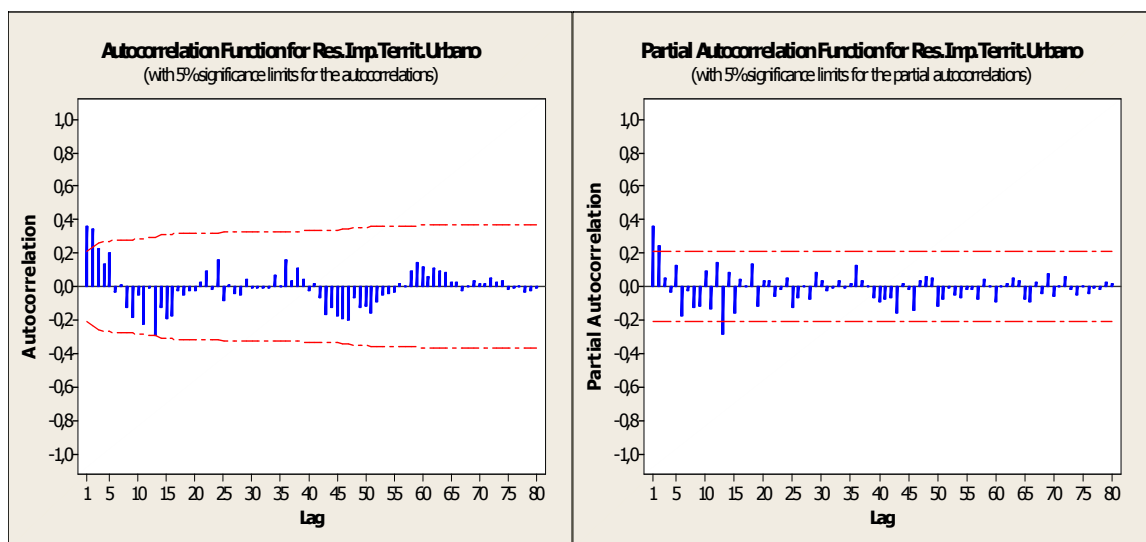


Figura 20: FAC e FACP para os Resíduos do ajuste de Regressão série Log 10 do Imposto Territorial Urbano

A Figura 20 apresenta a Função de Correlação e Função de Correlação Parcial e nos revela, a princípio, tratar-se de um processo de médias móveis, demonstrando a dependência temporal entre os resíduos.

Considerando que o pressuposto de independência dos resíduos foi violado, inserimos mais uma variável preditora no modelo de regressão. Essa variável que representa os resíduos oriundos da estimativa inicial é dada pela seguinte fórmula:

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

O teste de Durbin Watson se restringe a correlações de primeira ordem adotamos a defasagem de um passo atrás. Esta proposta evidencia que o valor observado em  $t_{(0)}$  guarda correlação com o valor observado no tempo  $t_{(-1)}$ . Ou seja, o erro observado em  $t_{(-1)}$  será utilizado como variável preditora para o valor estimado em  $t_{(0)}$ .

### 3.3.2 – Ajuste do modelo de regressão com inserção do erro de previsão

Este modelo tem como variável resposta o Log10 do Imposto Territorial Urbano e como variáveis predictoras um contador que identifica a sequência da série temporal (tempo), o fato de ser a previsão relativa ao mês de janeiro ou não (variável jan), a variável *dummy* “jan” multiplicada pelo “tempo” (janXtempo), a variável *dummy* novembro de 2006 (Nov\_2006) e a variável resíduo $_{(t-1)}$ .

#### Equação de regressão:

$$\text{Log10 Imp.Territ.Urbano} = 6,10 - 0,215(\text{Nov\_2006}) + 0,00274(\text{tempo}) + 1,23(\text{jan}) \\ - 0,00215(\text{janXtempo}) + 0,390 \text{ resíduo }_{(t-1)}$$

Tabela 5 – Estimativas dos parâmetros da regressão para a série Log10 Imposto Territorial

| Variável    | Coefficiente | Desvio padrão | Teste T | P-valor |
|-------------|--------------|---------------|---------|---------|
| Constante   | 6,10001      | 0,00841       | 725,68  | 0,000   |
| Nov-2006    | -0,21457     | 0,03726       | -5,76   | 0,000   |
| Tempo       | 0,0027393    | 0,0001644     | 16,67   | 0,000   |
| Jan         | 1,23375      | 0,03292       | 37,47   | 0,000   |
| JanXtempo   | -0,0021455   | 0,0006121     | -3,51   | 0,001   |
| $e_{(t-1)}$ | 0,3900       | 0,1067        | 3,65    | 0,000   |

Considerando os dados da Tabela 5, constatamos que os parâmetros são significativos para o modelo, ao nível de significância de 5%. O modelo ajustado pela regressão linear simples apresenta um  $R^2$  de 98,8%;  $R^2$  ajustado de 98,7%; Valor do teste F de 1.313,79; p-valor igual a 0,000 e estatística de Durbin-Watson (d) 2,16405.

Utilizando a tabela de valores críticos da estatística de Durbin-Watson para  $\alpha=0,05$  e  $n= 88$  e  $k=4$ , temos:  $d_L = 1,57$  e  $d_U = 1,75$ .

Uma vez que  $(4 - d) = 1,8359 > d_u = 1,75$  e que  $d = 2,16405 > d_u = 1,75$ , não há evidências da existência de correlações positivas ou negativas entre os erros, assim, não é rejeitada a hipótese de independência. Podemos, pois, admitir que os erros são independentes.

Na validação do modelo pelos resíduos apresentados na Figura 21, podemos constatar que se tratam de resíduos homocedásticos, independentes e estão distribuídos normalmente, conforme teste de normalidade apresentado na Figura 22.

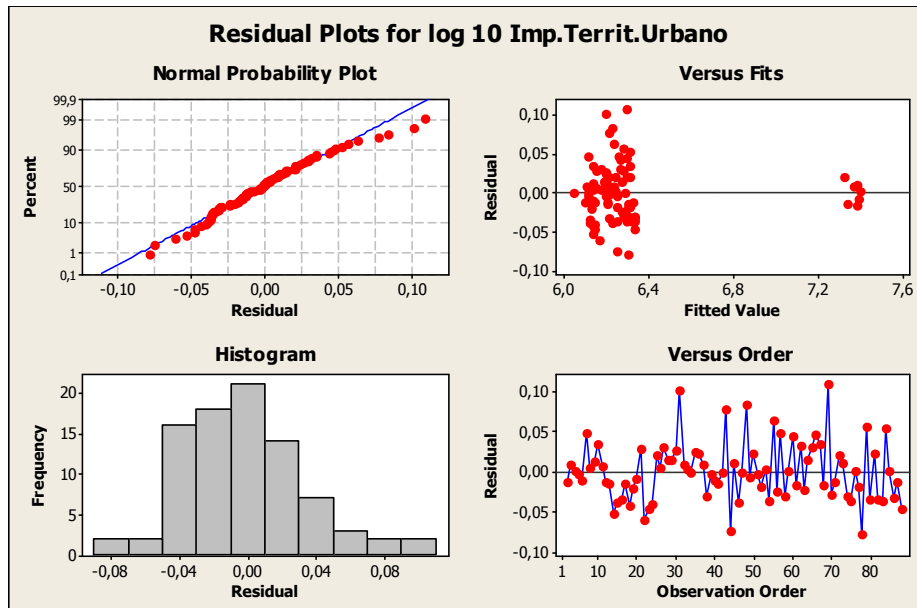


Figura 21: Resíduos do ajuste de regressão para a série Log10 do Imposto Territorial Urbano

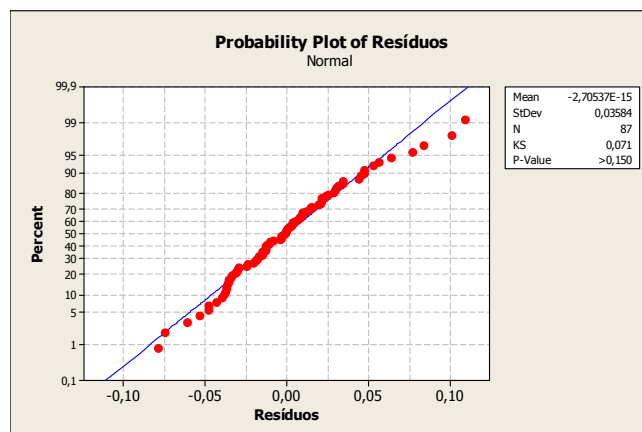


Figura 22: Teste de normalidade Resíduos do ajuste de regressão para a série Log10 do Imposto Territorial Urbano

### 3.3.3 - Previsão usando o modelo de regressão ajustado

Tabela 6: Previsões (1 passo a frente) maio a setembro de 2009 – Imposto Territorial, valores em Reais (R\$)

| <b>Mês</b> | <b>Previsão limite inferior</b> | <b>Previsão</b> | <b>Previsão limite superior</b> | <b>Valor observado</b> | <b>Erro previsão</b> |
|------------|---------------------------------|-----------------|---------------------------------|------------------------|----------------------|
| maio       | 1.872.940,83                    | 2.097.124,19    | 2.349.935,50                    | 1.998.461,09           | -98.663,10           |
| junho      | 1.967.814,32                    | 2.221.263,09    | 2.462.951,64                    | 2.117.016,00           | -104.247,09          |
| julho      | 1.990.042,10                    | 2.235.321,52    | 2.491.252,43                    | 2.407.044,38           | 171.722,86           |
| agosto     | 1.982.528,95                    | 2.249.468,94    | 2.482.335,41                    | 2.083.496,15           | -165.972,79          |
| setembro   | 2.031.028,13                    | 2.263.705,89    | 2.543.572,16                    | 2.143.504,29           | -120.201,60          |

Para fins de cálculo dos valores projetados, utilizamos o valor do erro registrado em abril/2009 para a estimativa do mês de maio/2009. Para os demais meses não foram utilizados valores de erro como variável preditora, posto não haver valores específicos e a média do erro ser igual a zero.

A Tabela 6 apresenta as estimativas um passo a frente para o período de Maio a Setembro de 2009 relativas ao Log 10 do Imposto Territorial. Os valores em logaritmo foram ajustados para a moeda corrente. Em valores absolutos, o maior erro da estimativa se deu em julho/2009 e o menor em maio/2009. Em termos percentuais, os erros representam respectivamente 7,13% e 4,94%.

As previsões dos limites inferior e superior foram construídas observando-se um intervalo de confiança de 95%, conforme metodologia explicitada em MONTGOMERY e RUNGER (2003, p.246).

### 3.4 - ANÁLISE DESCRITIVA DOS DADOS - IMPOSTO SOBRE TRANSMISSÃO DE BENS IMÓVEIS (ITBI)

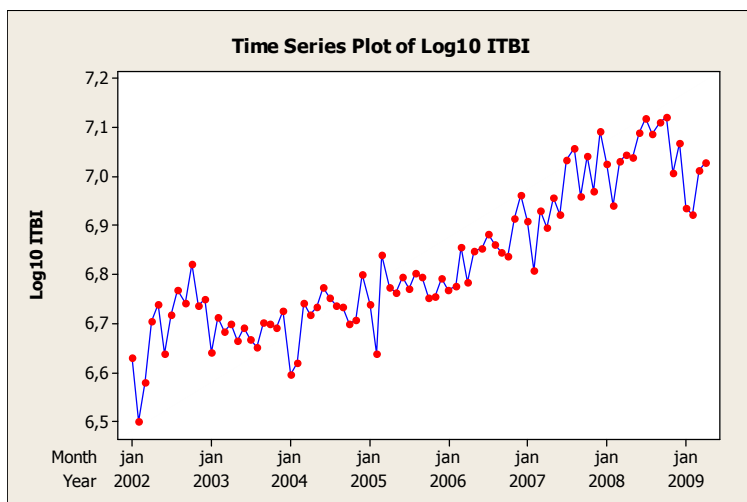


Figura 23: Gráfico da série Log10 ITBI

A Figura 23 apresenta o gráfico da série Imposto sobre Transmissão de Bens Imóveis (ITBI) para o período de janeiro de 2002 a abril de 2009. Observamos que a série apresenta tendência de crescimento ao longo de todo o período analisado.

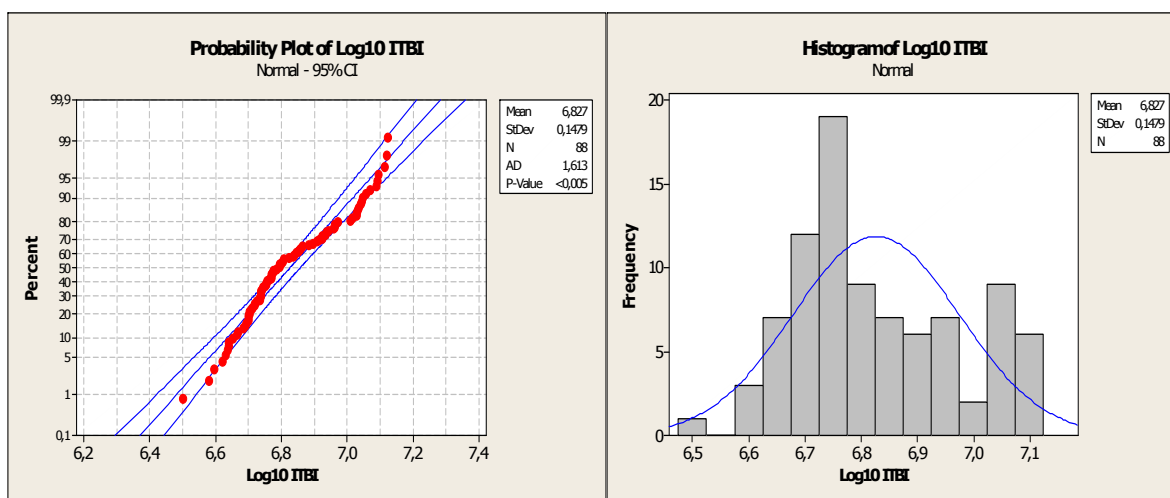


Figura 24: Gráfico de probabilidade e histograma da série Log10 ITBI

A análise dos gráficos da Figura 24 evidencia que os dados da série Imposto sobre Transmissão de Bens Imóveis – ITBI não se aproximam de uma Distribuição Normal, já que o p-valor apresentado é menor que 0,05.

### 3.4.1 - Ajuste do modelo de Alisamento Exponencial Holt Winters

O conjunto de dados estudado, conforme se percebe na análise descritiva, apresenta tendência de crescimento e não estacionariedade. Após testes dos modelos de alisamento exponencial, detectamos que o modelo de alisamento Holt Winters com sazonalidade aditiva foi o que produziu o melhor ajuste, pois apresentou em relação aos demais os menores erros.

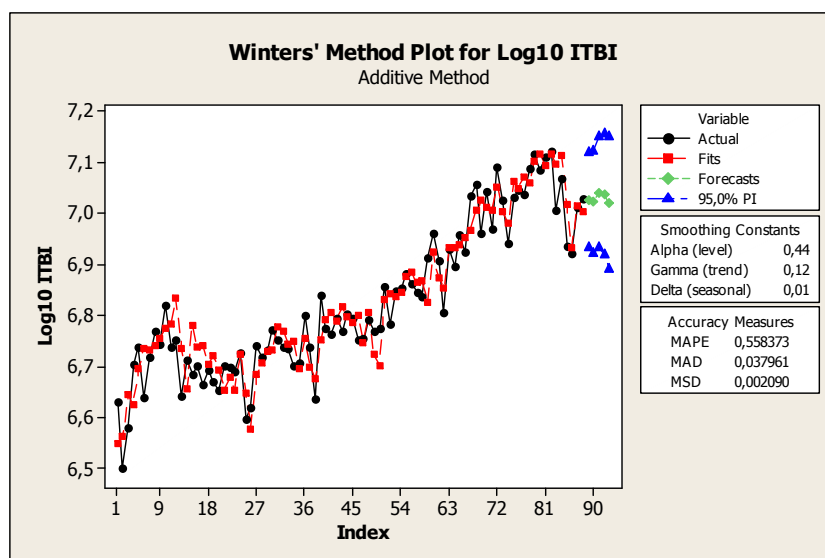


Figura 25: Alisamento Exponencial Holt Winters para a série Log10 ITBI

A Figura 25 apresenta a modelagem da série do Imposto sobre Transmissão de Bens Móveis - ITBI através do alisamento exponencial Holt Winters com tendência sazonal aditiva. Podemos observar que o modelo utilizando acompanha a tendência de crescimento apresentada pela série

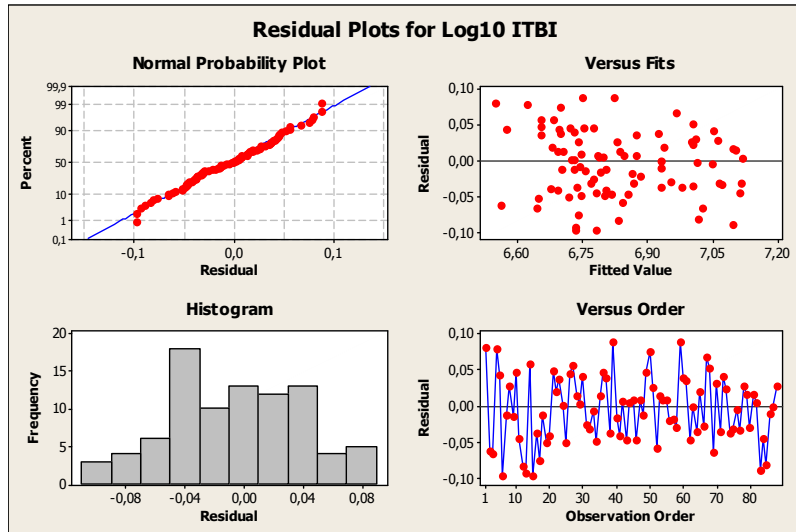


Figura 26: Resíduos do Alisamento Exponencial Holt Winters para a série Log10 ITBI

A modelagem através do método de alisamento exponencial não pressupõe normalidade dos resíduos. Contudo, foi gerado o gráfico da Figura 26 que apresenta os resíduos da modelagem da série e demonstra que os resíduos são homocedásticos e normais, conforme se depreende da análise do gráfico da Figura 27.

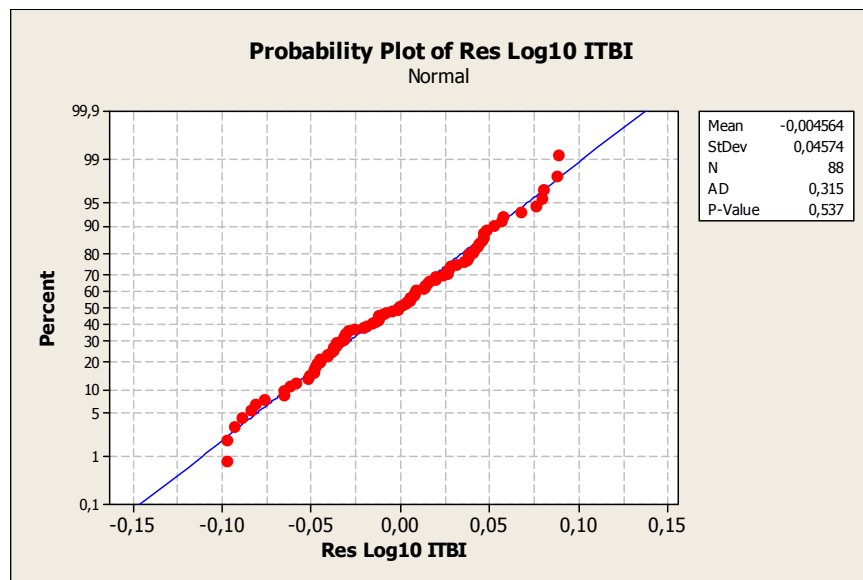


Figura 27: Teste de normalidade para os Resíduos da série Log10 ITBI

A Figura 28 apresenta a Função de Correlação e Função de Correlação Parcial dos resíduos do ajuste da série do Imposto sobre Transmissão de Bens Imóveis. Os resíduos se comportam como um ruído branco. Não é necessário ajustar um modelo para os resíduos.

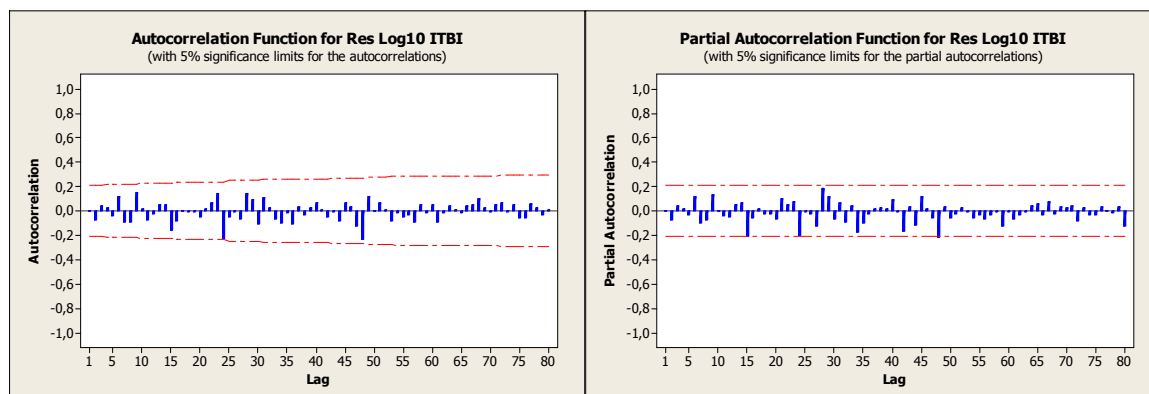


Figura 28: FAC e FACP para os Resíduos do ajuste por alisamento Holt Winters para da série Log10 ITBI

### 3.4.2 – Previsão usando o modelo ajustado

Tabela 7: Previsões (1 passo a frente) maio a setembro de 2009 – Imposto sobre Transmissão de Bens Imóveis – ITBI, valores em Reais (R\$)

| Mês      | Previsão<br>Limite Inferior | Previsão      | Previsão<br>Limite Superior | Valor<br>observado | Erro          |
|----------|-----------------------------|---------------|-----------------------------|--------------------|---------------|
| maio     | 8.598.927,94                | 10.652.462,16 | 13.196.237,26               | 12.287.930,93      | -1.635.468,77 |
| junho    | 8.367.665,56                | 10.541.684,16 | 13.280.370,16               | 14.033.259,16      | -3.491.575,00 |
| julho    | 8.563.125,50                | 11.006.519,31 | 14.147.375,79               | 14.445.704,66      | -3.439.185,35 |
| agosto   | 8.302.655,22                | 10.915.911,35 | 14.351.994,06               | 14.758.064,03      | -3.842.152,68 |
| setembro | 7.774.091,85                | 10.475.626,36 | 14.115.854,37               | 13.736.728,65      | -3.261.102,29 |

A Tabela 7 apresenta os valores reais, observados no período de maio a setembro de 2009 relativos ao Imposto sobre Transmissão de Bens Imóveis e as estimativas um passo a frente, utilizando o modelo alisamento exponencial de Holt Winters. Os limites inferior e superior das previsões foram calculados automaticamente pelo algoritmo do aplicativo minitab, observando um intervalo de confiança de 95%.

Após conversão dos valores em moeda, observamos que em termos absolutos o maior erro de previsão observado se deu em agosto/2009 e o menor em maio/2009. Em termos percentuais representam respectivamente 26,03% e 13,31%.

O erro apresentado é muito significativo em relação ao que foi observado no período. Uma possível explicação para este fenômeno é a ocorrência de um aumento significativo nas transações comerciais relativas aos imóveis na capital que não foi capturado pelo modelo utilizado. A cidade vive uma expansão imobiliária motivada pelas construções de infra-estrutura de transporte, construção da nova sede do governo estadual, desapropriações para construção de equipamentos públicos, dentre outros, que, além de aumentar o número de operações comerciais, opera um aumento no preço dos imóveis. Qualquer modelo seria contaminado por esta variável dificultando a projeção de valores, visto que a serie histórica coletada se refere a um período anterior a este fenômeno, ou no qual sua repercussão sobre a série ainda não era tão visível.

### 3.5 - ANÁLISE DESCRITIVA DOS DADOS – IMPOSTO SOBRE SERVIÇOS (ISS)

A Figura 29 apresenta a série LOG10 ISS no período de janeiro de 2002 a abril de 2009. vejamos que a série apresenta crescimento ao longo do tempo. Observa-se que os valores recolhidos mensalmente vêm aumentando significativamente ao longo dos anos estudados, evidenciando tendência de crescimento da série.

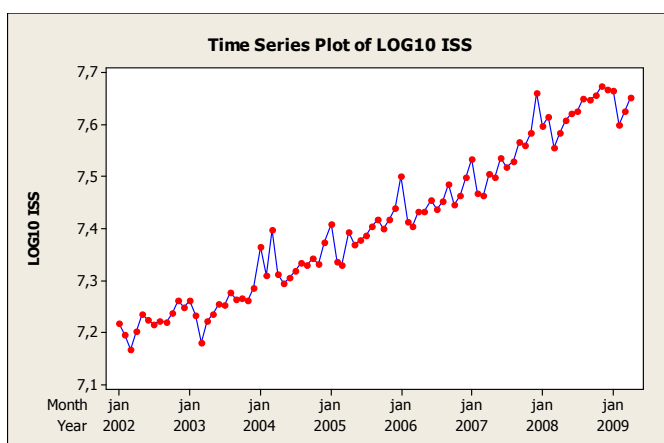


Figura 29: Gráfico da série Log10 ISS

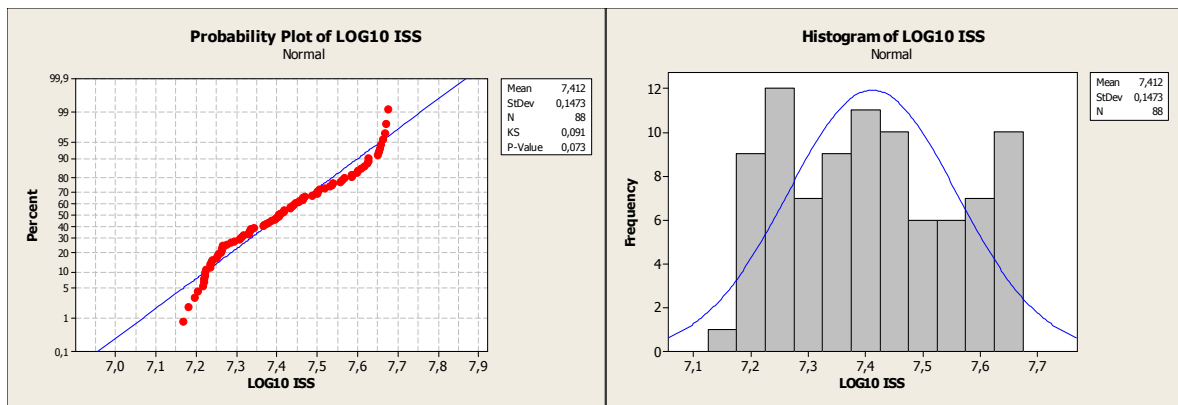


Figura 30: Gráfico de probabilidade e histograma da série Log10 ISS

A análise dos gráficos da Figura 30 evidencia que os dados da série Log10 ISS tem distribuição normal, já que o p-valor apresentado é maior que o nível de significância de 0,05.

### 3.5.1 - Ajuste do modelo de Alisamento Exponencial Holt Winters

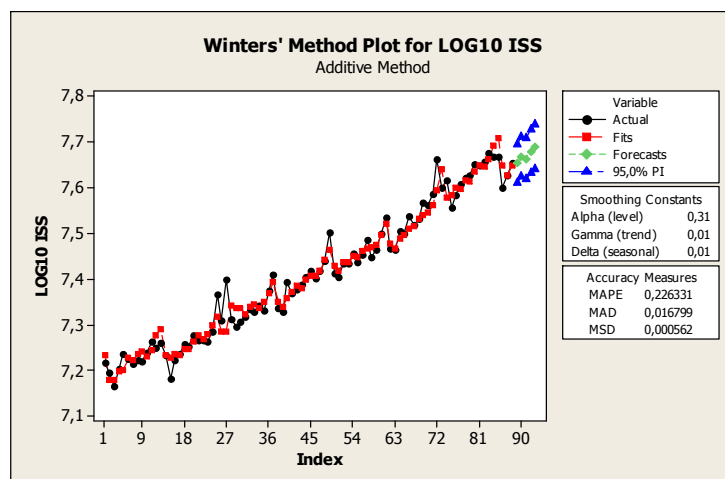


Figura 31: Alisamento Exponencial Holt Winters para a série LOG10 ISS

A Figura 31 apresenta a modelagem da série Log10 ISS através do método Holt Winters Aditivo e a Figura 32 os resultados da modelagem. Podemos observar que o modelo utilizado acompanha a tendência de crescimento apresentado pela série.

Outros modelos de alisamento foram testados, porém o modelo aqui utilizado foi o que apresentou as menores medidas de erro (*accuracy measures*)

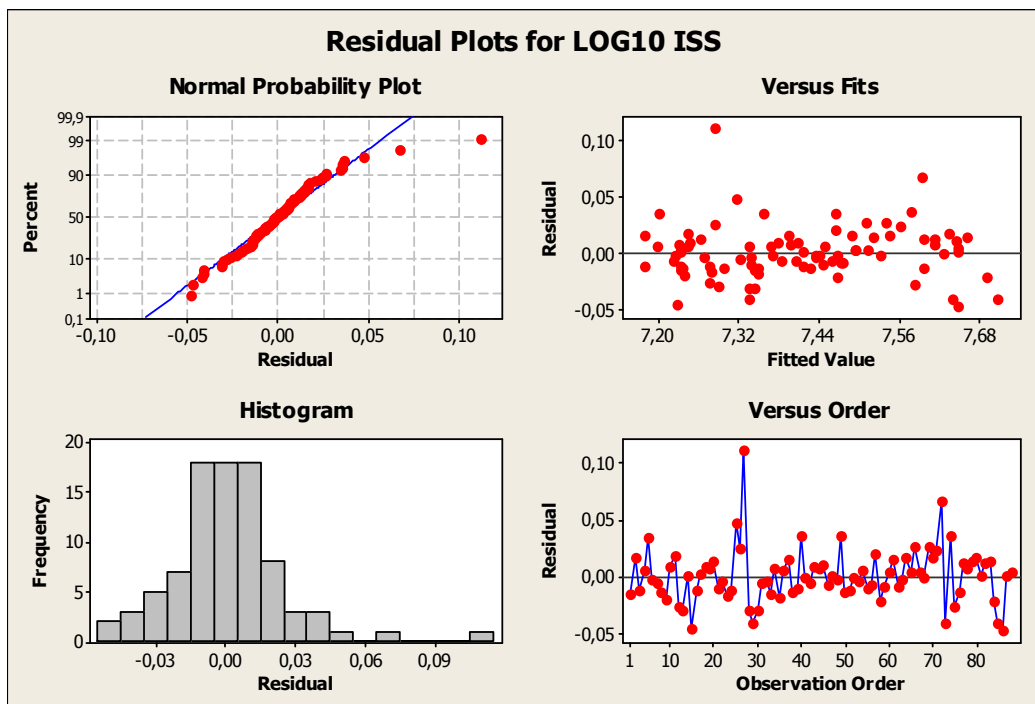


Figura 32: Resíduos do Alisamento Exponencial Holt Winters para a série Log10 ISS

A modelagem através do método de alisamento exponencial não pressupõe normalidade dos resíduos, dispensando o teste de normalidade.

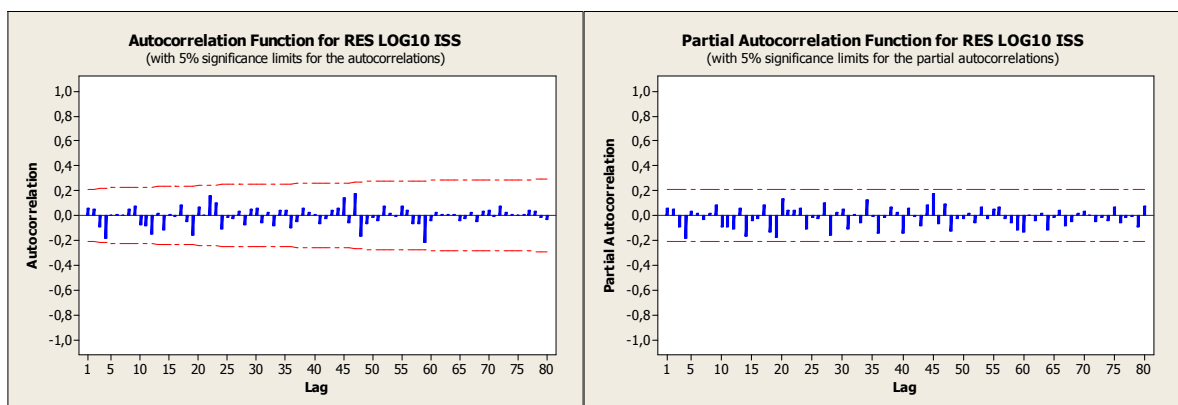


Figura 33: FAC e FACP para os resíduos do ajuste por alisamento exponencial Holt Winters para a série Log10 ISS

A Figura 33 apresenta a Função de Correlação e Função de Correlação Parcial dos resíduos do ajuste da série do Log10 ISS. Os resíduos do ajuste resultaram num ruído branco.

### 3.5.2 - Previsão usando o modelo ajustado

Tabela 8: Previsões (1 passo a frente) maio a setembro de 2009 – ISS, valores em Reais (R\$)

| <b>Mês</b> | <b>Previsão<br/>Limite Inferior</b> | <b>Previsão</b> | <b>Previsão<br/>Limite Superior</b> | <b>Valor<br/>observado</b> | <b>Erro</b>    |
|------------|-------------------------------------|-----------------|-------------------------------------|----------------------------|----------------|
| maio       | 40.872.246,80                       | 44.935.543,61   | 49.401.888,67                       | 43.586.176,62              | 1.349.366,99   |
| junho      | 42.093.747,75                       | 46.446.179,90   | 51.249.284,39                       | 45.064.658,05              | 1.381.521,85   |
| julho      | 41.531.639,02                       | 46.019.299,12   | 50.992.692,23                       | 46.343.708,12              | - 324.409,00   |
| agosto     | 42.944.861,52                       | 47.811.239,18   | 53.228.345,93                       | 50.279.337,18              | - 2.468.098,00 |
| setembro   | 43.575.910,96                       | 48.765.198,91   | 54.571.944,34                       | 46.834.611,78              | 1.930.587,13   |

A Tabela 8 apresenta as estimativas um passo a frente para o período de Maio a Setembro de 2009 relativas ao Log 10 do ISS, utilizando o modelo de alisamento exponencial Holt Winters. Após conversão dos valores em moeda, observamos que em termos absolutos o maior erro de previsão observado se deu em agosto/2009 e o menor em julho/2009. Em termos percentuais representam respectivamente 4,91% e 0,70%.

Os limites inferior e superior das previsões foram calculados automaticamente pelo algoritmo do aplicativo minitab, observando um intervalo de confiança de 95%.

### 3.6 ANÁLISE DESCRITIVA DOS DADOS – ISS AUTONOMOS E PROFISSIONAIS LIBERAIS

A Figura 34 apresenta o gráfico da série Log 10 ISS Autônomos e Profissionais Liberais para o período de janeiro de 2002 a abril de 2009. Observamos que a série aparenta ser estacionária, com grande variabilidade entre os recolhimentos nos meses que compõem, evidenciado sazonalidade que precisa ser estudada.

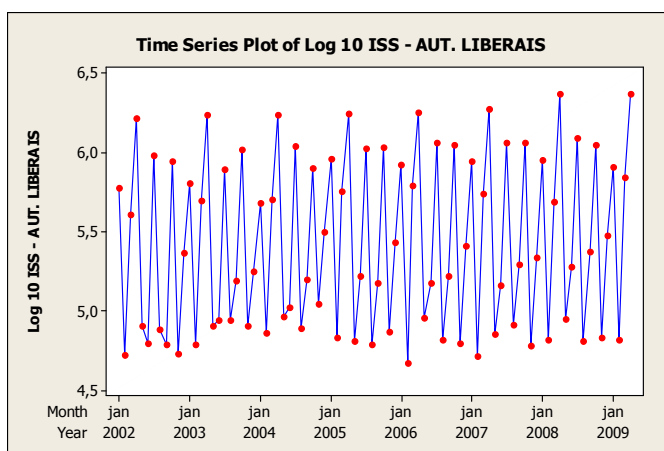


Figura 34: Gráfico da série Log10 ISS Autônomos e Liberais

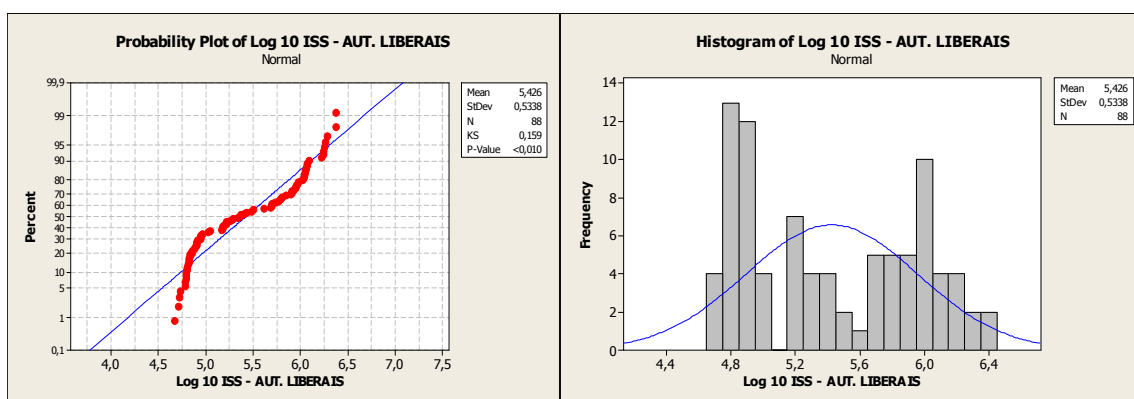


Figura 35: Gráfico de probabilidade e histograma da série Log10 ISS Autônomos e Liberais

A análise dos gráficos da Figura 35 evidencia que os dados da série Log10 ISS Autônomos e Liberais não se aproximam de uma Distribuição Normal, já que o p-valor apresentado é menor que 0,05.

### 3.6.1 - Ajuste do modelo de regressão

Para o caso das receitas municipais com o recolhimento do ISS de profissionais Autônomos e Liberais (série Log10 ISS Autônomos e Liberais), optamos por ajustar um modelo de regressão. Este modelo tem como variável resposta o Log10 do ISS Autônomos e Liberais e como variáveis preditoras um contador que identifica a sequência da série temporal (tempo), o fato de ser a previsão relativa aos meses do ano, para isso utilizamos variáveis *Dummy*, para cada mês do ano.

#### Equação de regressão:

$$\text{Log10 ISS Aut. Liberais} = 5,34 + 0,00178(\text{tempo}) + 0,459(\text{Jan}) - 0,636(\text{Fev}) + 0,314(\text{Mar}) + 0,861(\text{Abr}) - 0,502(\text{Mai}) - 0,324(\text{Jun}) + 0,612(\text{Jul}) - 0,549(\text{Ago}) - 0,237(\text{Set}) + 0,593(\text{Out}) - 0,568(\text{Nov})$$

Tabela 9 – Estimativas dos parâmetros de regressão para a série Log10 ISS autônomos e Liberais

| Variável  | Coefficiente | Desvio padrão | Teste T | P-valor |
|-----------|--------------|---------------|---------|---------|
| Constante | 5,33539      | 0,01881       | 283,63  | 0,000   |
| Tempo     | 0,0017752    | 0,0003672     | 4,83    | 0,000   |
| Jan       | 0,45913      | 0,02967       | 15,48   | 0,000   |
| Fev       | -0,63558     | 0,02966       | -21,43  | 0,000   |
| Mar       | 0,31411      | 0,02966       | 10,59   | 0,000   |
| Abr       | 0,86137      | 0,02967       | 29,04   | 0,000   |
| Mai       | -0,50242     | 0,03154       | -15,93  | 0,000   |
| Jun       | -0,32396     | 0,03153       | -10,28  | 0,000   |
| Jul       | 0,61218      | 0,03152       | 19,42   | 0,000   |
| Ago       | -0,54926     | 0,03151       | -17,43  | 0,000   |
| Set       | -0,23708     | 0,03151       | -7,52   | 0,000   |
| Out       | 0,59278      | 0,03152       | 18,81   | 0,000   |
| Nov       | -0,56798     | 0,03153       | -18,02  | 0,000   |

Considerando os dados da Tabela 9, constatamos que os parâmetros são significativos para o modelo, ao nível de significância de 5%.

O modelo ajustado pela regressão linear simples apresenta um  $R^2$  de 97,7%;  $R^2$  ajustado de 97,3%; Valor do teste F de 265,20; p-valor igual a 0,000 e estatística de Durbin-Watson (d) 2,22376.

Utilizando a tabela de valores críticos da estatística de Durbin-Watson para  $\alpha=0,05$ ,  $n= 88$  e  $k=12$ , temos:  $d_L = 1,52$  e  $d_U = 1,77$ .

Uma vez que  $(4 - d) = 1,77624 > d_U = 1,77$  e que  $d = 2,22376 > d_U = 1,77$ , não há evidências da existência de correlações positivas ou negativas entre os erros, assim, não é rejeitada a hipótese de independência. Podemos, pois, admitir que os erros são independentes.

Para testar a validade do modelo, analisamos os resíduos apresentados na regressão, que são mostrados na Figura 36.

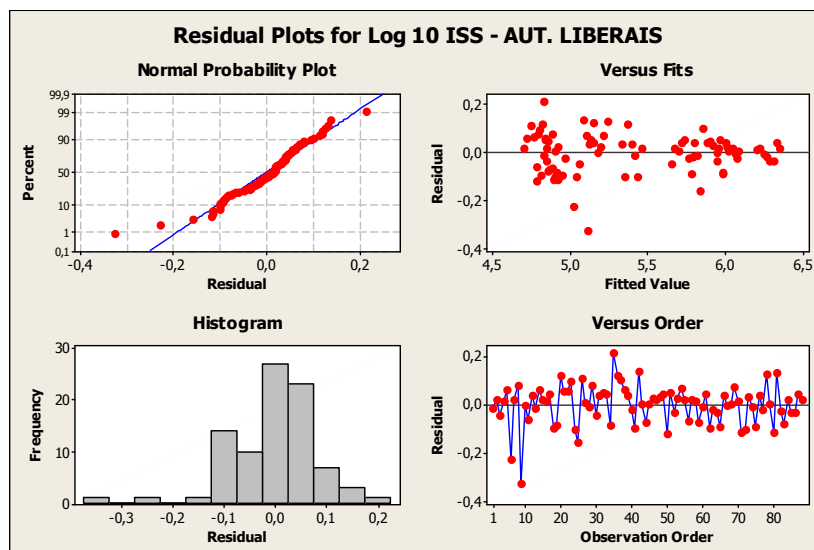


Figura 36: Resíduos do ajuste de Regressão para a série Log10 ISS Autônomos e Liberais

Na validação do modelo pelos resíduos apresentados na Figura 36, podemos inferir que se tratam de resíduos homocedásticos, independentes e estão distribuídos normalmente.

A Figura 37 apresenta o teste de normalidade dos resíduos da regressão para o Log10 ISS – autônomos e Liberais demonstrando que estes estão distribuídos conforme a distribuição normal, p-valor 0,074.

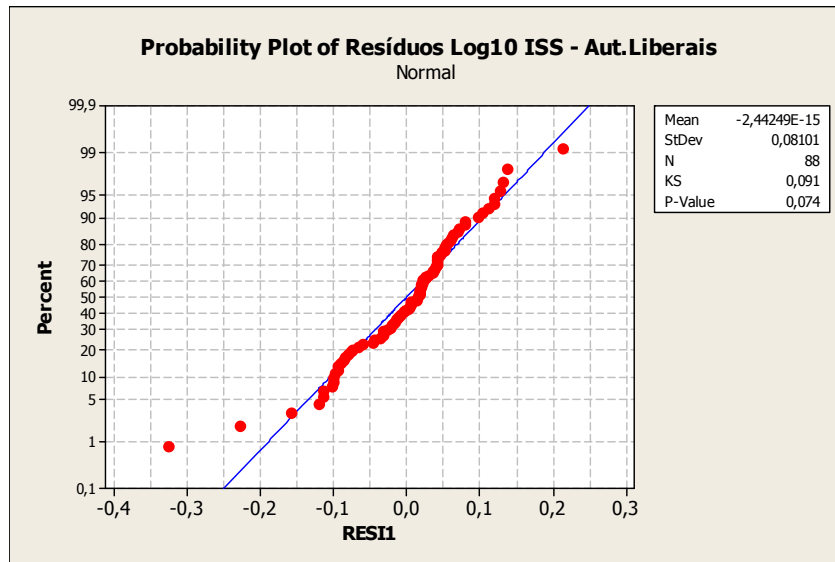


Figura 37: Teste de Normalidade para os Resíduos do ajuste de Regressão para o Log10 ISS – Autônomos e Liberais

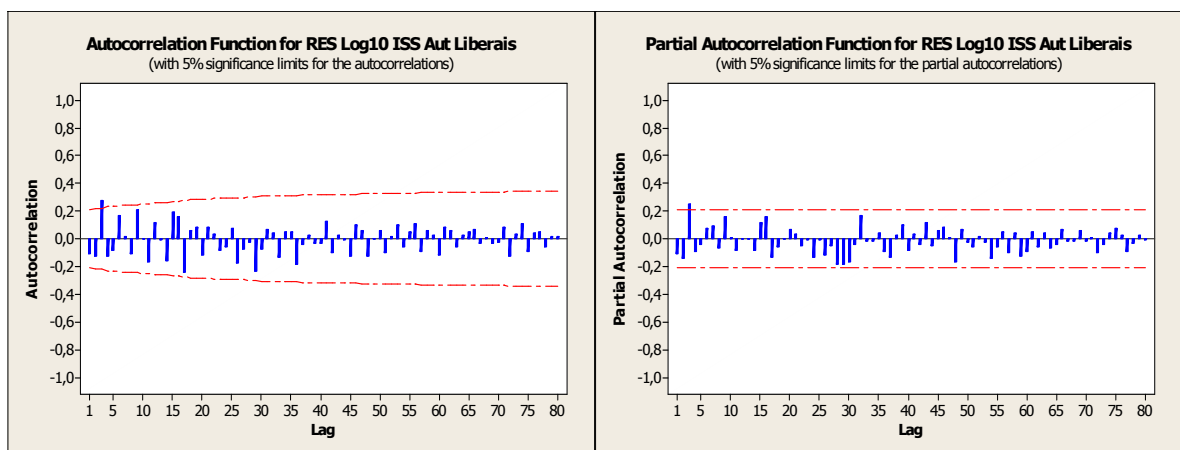


Figura 38: FAC e FACP para os Resíduos do ajuste de regressão para Log10 ISS Autônomos e Liberais

A Figura 38 que apresenta a Função de Correlação e Função de Correlação Parcial dos resíduos do ajuste de Regressão para Log10 ISS – Aut.Liberais nos revela que trata-se de um processo de médias móveis – MA(3).

Os pressupostos para validação do modelo não foram violados, concluímos pela adequação do modelo de regressão para o Log10 ISS – Aut.Liberais. Porém, para melhorar o desempenho do modelo utilizamos os erros de previsão ( $e_t = y_t - \hat{y}_t$ ) do modelo de regressão anterior como variável preditora de um novo modelo.

Posto que os erros de estimação estão correlacionados, descrevendo um processo de médias móveis (MA3) – conforme se verifica na FAC e FACP da Figura xxx – adotamos uma defasagem de três passos atrás para definir essa nova variável preditora. Nestes termos, o valor observado em  $t_{(0)}$  está correlacionado ao valor observado no tempo  $t_{(-3)}$ . Ou seja, o erro observando em  $t_{(-3)}$  será utilizado como variável preditora para o valor estimado em  $t_{(0)}$ .

### 3.6.2 – Ajuste do modelo de regressão com inserção do erro de previsão

Este modelo tem como variável resposta o Log10 do ISS Autônomos e Liberais e como variáveis predictoras um contador que identifica a sequência da série temporal (tempo), o fato de ser a previsão relativa ao meses do ano, para isso utilizamos variáveis *Dummy*, para cada mês do ano e a variável resíduo $_{(t-3)}$ .

#### Equação de regressão:

$$\begin{aligned} \text{Log10 ISS Aut. Liberais} = & 5,3353 + 0,0017420(\text{tempo}) + 0,46083(\text{Jan}) \\ & - 0,63881 (\text{Fev}) + 0,32036(\text{Mar}) + 0,86076(\text{Abr}) - 0,50447(\text{Mai}) \\ & - 0,32303(\text{Jun}) + 0,61225(\text{Jul}) - 0,554994(\text{Ago}) - 0,23772(\text{Set}) \\ & + 0,59217(\text{Out}) - 0,56856(\text{Nov}) + 0,2802 \text{ resíduo } (t-3) \end{aligned}$$

Tabela 10 – Estimativas dos parâmetros da regressão para a série Log10 ISS Autônomos e Liberais

| Variável      | Coefficiente | Desvio padrão | Teste T | P-valor |
|---------------|--------------|---------------|---------|---------|
| Constante     | 5,33753      | 0,01990       | 268,19  | 0,000   |
| Tempo         | 0,0017420    | 0,0003823     | 4,56    | 0,000   |
| Jan           | 0,46083      | 0,03107       | -20,55  | 0,000   |
| Fev           | -0,63881     | 0,03108       | 10,95   | 0,000   |
| Mar           | 0,32036      | 0,03110       | 10,30   | 0,000   |
| Abr           | 0,86076      | 0,02923       | 29,45   | 0,000   |
| Mai           | -0,50447     | 0,031111      | -16,22  | 0,000   |
| Jun           | -0,32303     | 0,03109       | -10,39  | 0,000   |
| Jul           | 0,61225      | 0,03107       | 19,71   | 0,000   |
| Ago           | -0,54994     | 0,03105       | -17,71  | 0,000   |
| Set           | -0,23772     | 0,03046       | -7,66   | 0,000   |
| Out           | 0,59217      | 0,03104       | 19,07   | 0,000   |
| Nov           | -0,56856     | 0,03105       | -18,31  | 0,000   |
| Resíduo (t-3) | 0,2802       | 0,1140        | 2,46    | 0,016   |

Considerando os dados da Tabela 10, constatamos que os parâmetros são significativos para o modelo, ao nível de significância de 5%.

O modelo ajustado pela regressão linear simples apresenta um  $R^2$  de 97,8%;  $R^2$  ajustado de 97,4%; Valor do teste F de 246,61; p-valor igual a 0,000 e estatística de Durbin-Watson (d) 2,10030.

Utilizando a tabela de valores críticos da estatística de Durbin-Watson para  $\alpha=0,05$  e  $n=88$  e  $k=13$ , temos:  $d_L= 1,52$  e  $d_U= 1,77$ .

Uma vez que  $(4 - d) = 1,8997 > d_U = 1,77$  e que  $d = 2,10030 > d_U = 1,77$ , não há evidências da existência de correlações positivas ou negativas entre os erros, assim, não é rejeitada a hipótese de independência. Podemos, pois, admitir que os erros são independentes.

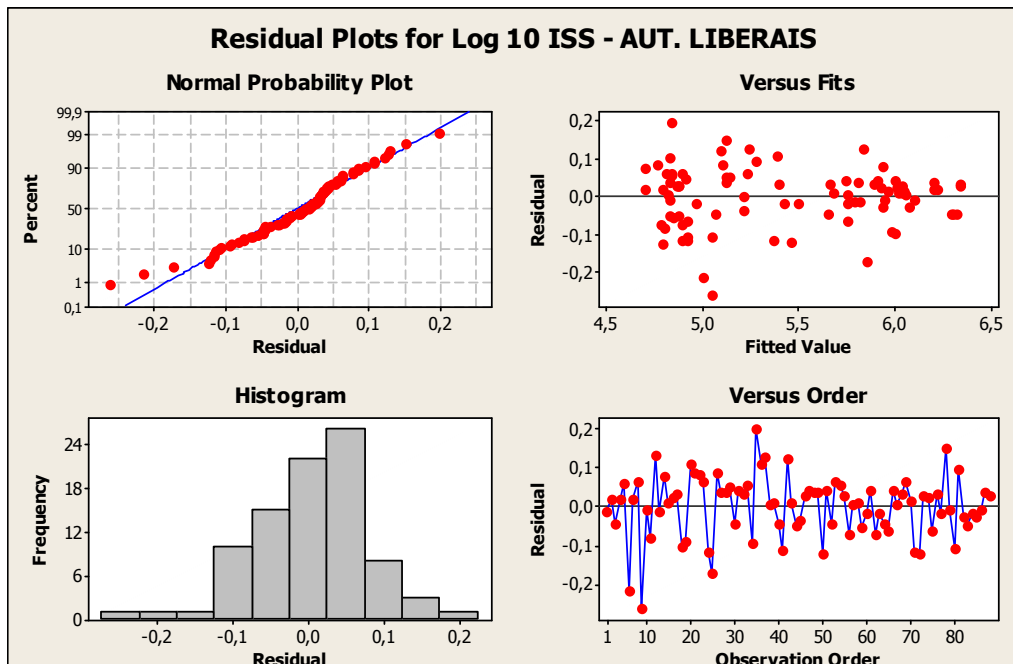


Figura 39: Resíduos do ajuste de regressão para o Log10 ISS Autônomos e Liberais

Na validação do modelo pelos resíduos apresentados na Figura 39, podemos inferir que se tratam de resíduos homocedásticos, independentes e estão distribuídos normalmente, conforme teste de normalidade apresentado na Figura 40.

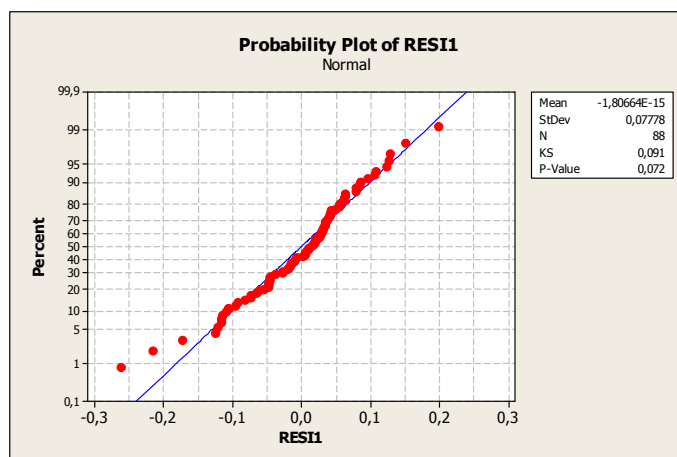


Figura 40: Teste de normalidade para os resíduos do ajuste de regressão para o Log10 ISS Autônomos e Liberais

### 3.6.3 - Previsão usando o modelo ajustado

Para fins de cálculo dos valores projetados, utilizamos o valor do erro registrado em fevereiro/2009 para a estimativa do mês de maio/2009, o erro de março/2009 para a estimativa do mês de junho/2009 e o erro de abril/2009 para a estimativa do mês de julho/2009. Para os demais meses não foram utilizados valores de erro como variável preditora, posto não haver valores específicos e a média do erro ser igual a zero.

Tabela 11: Previsões (1 passo a frente) maio a setembro de 2009 – Log10 ISS Autônomos e Liberais, valores em Reais (R\$)

| <b>Mês</b> | <b>Previsão limite inferior</b> | <b>Previsão (R\$)</b> | <b>Previsão limite superior</b> | <b>Valor observado</b> | <b>Erro previsão</b> |
|------------|---------------------------------|-----------------------|---------------------------------|------------------------|----------------------|
| maio       | 61.971,14                       | 95.306,89             | 146.574,53                      | 114.561,12             | 19.254,23            |
| junho      | 99.058,18                       | 152.399,15            | 234.378,35                      | 233.256,73             | 80.857,58            |
| julho      | 845.050,20                      | 1.299.326,41          | 1.998.262,30                    | 1.324.381,27           | 25.054,86            |
| agosto     | 57.689,09                       | 88.685,78             | 136.391,87                      | 88.527,82              | -157,96              |
| setembro   | 118.865,45                      | 182.732,59            | 281.028,58                      | 207.421,27             | 24.688,68            |

A Tabela 11 apresenta as estimativas um passo a frente para o período de Maio a Setembro de 2009 relativas ao Log10 ISS – Aut. Liberais, utilizando o modelo de regressão. Os valores em logaritmo foram ajustados para a moeda corrente. Em valores absolutos, o maior erro da estimativa se deu em junho/2009 e o menor em agosto/2009. Em termos percentuais, os erros representam respectivamente 34,66% e 0,18%.

As previsões dos limites inferior e superior foram construídas observando-se um intervalo de confiança de 95%, conforme metodologia explicitada em MONTGOMERY e RUNGER (2003, p.246).

## 4 DISCUSSÕES E CONCLUSÃO

Obras, serviços de saúde, de segurança e educação, dentre outros, são ofertados e mantidos pelo poder público municipal e interferem no bem estar da população e só se realizam mediante a existência de recursos financeiros. Sem receitas financeiras não há como realizar tais atividades.

O planejamento da vida financeira do município é essencial para concretizar os objetivos de um governo. Neste sentido, previamente à execução e oferta de serviços públicos, devem ser confeccionados e aprovados orçamentos fiscais que demonstrem a capacidade de pagamento e de investimento.

O poder público prescinde de instrumentos de previsão de receitas capazes de alimentar um processo de planejamento do qual resulte ações que beneficiem o cidadão contribuinte dos impostos que se arrecadam. Para isso a análise das séries temporais relativas aos impostos próprios é fundamental neste processo de planejamento.

Um problema comum para a constituição dos orçamentos de receitas se relaciona ao erro que podem ser verificados durante o processo de previsão.

Vejamos que os erros de previsão negativos, ou seja, os casos em que o valor observado é menor que o valor estimado, acarretam problemas. Quando os valores estimados se mostram maiores que os reais, significa que o setor público está propenso a incorrer em déficit orçamentário, caso não tenha tomado os devidos cuidados com o controle de gastos, posto que a receita esperada e as despesas estarão em descompasso. No sentido oposto, estimativas cujo valor é inferior ao valor observado são a princípio mais atrativas, pela possibilidade da ocorrência de superávit orçamentário, preconizado pela Lei de Responsabilidade Fiscal.

Assim, as informações produzidas por ferramentas de previsão em séries temporais são utilizadas pelo setor público para a confecção de orçamentos de entrada de recursos financeiros, visando evitar erros orçamentários e suas implicações. Portanto, a escolha da ferramenta de previsão utilizada impacta no bom desempenho das atividades públicas.

Para algumas séries optamos por utilizar o modelo de regressão linear. Os resultados demonstraram que os modelos lineares apresentam alta qualidade de ajuste, como demonstrado pela estatística “R<sup>2</sup>” ajustado e pelos resultados da previsão. Em outras séries, utilizamos a técnica de alisamento exponencial para modelar e prever os valores futuros. Os resultados também foram satisfatórios, conforme análises descritas no corpo do texto.

Os modelos ARIMA foram utilizados para análise alternativa dos dados e comparação entre modelos. Todas as séries foram modeladas e previstas com utilização desta técnica. Os resultados dessa modelagem não foram citados expressamente, dado o volume de informação gerado e o necessário enxugamento do conteúdo deste trabalho. Acrescentamos que a técnica de modelagem ARIMA foi útil na formação de nossa convicção sobre a qualidade dos modelos escolhidos para ajuste final das séries e apresentação neste trabalho.

Devido às características das receitas municipais, é necessário um acompanhamento sistemático de previsão com validação de informações mais recentes, haja vista que os métodos de previsão demonstraram possuir boa modelagem dos dados, mas com uma capacidade de previsão que precisa ser melhorada em sua acurácia. Noutro momento, sugere-se a inclusão de uma nova variável preditora auto-regressiva nos modelos deficientes em previsão, a fim de melhorar a eficiência preditora.

A percebida dificuldade de fazer previsões decorre da sensibilidade das séries estudadas às mudanças no ambiente em que se dão as relações comerciais sujeitas a incidência dos impostos. O país, por consequência os municípios, foi prejudicado pela crise econômica que afetou as economias mundiais em 2007/2008, cujos efeitos ainda são perceptíveis. Porém, neste último ano de 2009, o país retomou o processo de crescimento econômico. Esse processo de retomada do crescimento econômico não foi considerado como variável nos modelos de previsão utilizados neste trabalho. Daí resultou um desempenho questionável no aspecto da previsibilidade dos modelos ajustados.

Contudo, é inegável que o estudo sistemático das receitas financeiras do município, utilizando modelos de previsão em séries temporais, é fundamental para o exercício das atividades inerentes ao governo.

## 5 BIBLIOGRAFIA

AKAIKE, H. Information theory as an extension of the maximum likelihood principle. In: INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY, 2, Budapest, 1973. Proceedings, Budapest, Akademia Kiadó, 1973, p.267-281.

BOWERMAN, B. L., O'CONNELL, R. T. *Forecasting and Time Series: An Applied Approach*, 3rd. ed. Wadsworth, Inc., 1993.

BOX, G. E. P.; JENKINS, G. M. *Time series analysis: forecasting and control*. San Francisco: Holden-Day, 1976.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. *Times series analysis: forecasting and control*. 3ª. Edição. São Francisco, Holden-Day, 1994.

BUENO, Rodrigo de Losso da Silveira. *Econometria de Séries Temporais*. São Paulo: Cengage Learning, 2008.

DURBIN, J., and WATSON, G. S. (1951) Testing for Serial Correlation in Least Squares Regression, II. *Biometrika* 38, 159–179.

KASMIER, Leonard. *Estatística Aplicada à Administração e Economia*. Tradução Adriano Silva Vale Cardoso. 4.ed. Porto alegre:Bookman, 2007.

MAKRIDAKIS, Spyros. WHEELWRIGTH, Steven C. *Forecasting Methods for Management*, 4ª Ed. New York, John Wiley e Sons Inc, 1985.

MONTGOMERY Douglas C., RUNGER George C., *Estatística Aplicada e Probabilidade para Engenheiros*, 2ª Edição, Ltc Editora, 2003.

MORETTIN, P. A.; TOLOI, C. M. C. *Análise de Séries Temporais* – São Paulo: Edgard Blucher, 2004.

MORETTIN, P. A.; TOLÓI, C. M. C. *Análise de séries temporais*. 2. ed. São Paulo: Edgard Blücher, 2006. - (Coleção Projeto Fisher)

SHIKIDA, Cláudio D. Econometria. IBMEC-MG Departamento de Economia. Versão 4. Julho, 2008.

SOUZA, R. C., CAMARGO, M. E. *Análise de Séries Temporais: Os Modelos ARIMA*, Rio de Janeiro: Gráfica e Editora Regional, 2004.

WERNER, L.; RIBEIRO, J. L. D. Demand forecasting: an application of the Box-Jenkins models in the technical assistance of personal computer. *Gest. Prod.* v.10, n.1, p.47-67, 2003.