



## PREVISÃO DE DEMANDA: UMA APLICAÇÃO DOS MODELOS BOX- JENKINS NA ÁREA DE ASSISTÊNCIA TÉCNICA DE COMPUTADORES PESSOAIS

**Liane Werner**

Departamento de Estatística, Universidade Federal do Rio Grande do Sul, Rua Bento Gonçalves, 9500, CEP 91509-000, Porto Alegre, RS, e-mail: lianew@mat.ufrgs.br

**José Luis Duarte Ribeiro**

Programa de Pós-graduação em Engenharia de Produção, Universidade Federal do Rio Grande do Sul, Praça Argentina, 9, 2º andar, sala LOPP, CEP 90040-020, Porto Alegre, RS, e-mail: ribeiro@ufrgs.br

### **Resumo**

---

*A previsão de demanda é uma atividade importante para auxiliar na determinação dos recursos necessários para a empresa. Neste artigo, a metodologia de Box-Jenkins foi utilizada para analisar dados históricos de uma empresa de assistência técnica de computadores pessoais e obter previsões do número de atendimentos. A empresa estudada apresenta três tipos de clientes diferenciados: contratos, garantia e avulsos. Como cada segmento de clientes tem suas peculiaridades, a previsão de demanda foi direcionada a cada tipo, buscando representar o comportamento de tendência e a sazonalidade por meio dos modelos de Box-Jenkins. A obtenção dos modelos mais adequados foi baseada na análise de gráficos e em testes estatísticos próprios da metodologia, os quais subsidiaram a decisão de adotar o modelo AR(1) para prever o número de atendimentos dos clientes tipo contrato, o modelo ARIMA(2,1,0) para os clientes tipo garantia e um modelo sazonal SARIMA(0,1,0)(0,1,1)<sub>12</sub> para os clientes tipo avulsos.*

**Palavras-chave:** *previsão de demanda, modelos Box-Jenkins, modelos ARIMA, séries temporais, assistência técnica.*

## 1. Introdução

Na última década, a globalização passou a ser a palavra-chave no mercado mundial. As empresas que desejam se manter nele, ou mesmo em seus mercados locais, precisam, antes de tudo, analisar a situação que as rodeiam e atentar aos rumos tomados no âmbito da globalização.

Segundo Makridakis *et al.* (1998), realizar previsões de demanda é importante para auxiliar na determinação dos recursos necessários para a empresa. Em tempos de abertura de mercados, essa atividade torna-se fundamental. Os mercados que podem ser acessados pela empresa, assim como a concorrência que os disputa, mudam continuamente, exigindo novas previsões de demanda em períodos mais curtos.

Na área de assistência técnica de computadores, prever a demanda dos serviços está intimamente ligado à disponibilização de recursos, tanto humanos como de equipamentos, para um atendimento eficaz aos clientes.

As previsões de demanda são elaboradas utilizando técnicas quanti e qualitativas ou, ainda, uma mistura de ambas. Segundo Pellegrini & Fogliatto (2000), métodos quantitativos utilizam dados históricos para prever a demanda em períodos futuros.

Uma das técnicas quantitativas mais difundida é a metodologia de Box-Jenkins, descrita por esses autores na década de 70. Os modelos de Box-Jenkins partem da idéia de que cada valor da série (temporal) pode ser explicado por valores prévios, a partir do uso da estrutura de correlação temporal que geralmente há entre os valores da série.

Segundo Abdel-Aal & Al-Garni (1997), os modelos Box-Jenkins têm sido largamente utilizados para modelagem e previsão em aplicações médicas, ambientais, financeiras e de engenharia. Os próprios autores aplicaram essa metodologia para prever o consumo mensal de energia elétrica no leste da Arábia Saudita.

Para Slini *et al.* (2001), a modelagem de séries temporais é apropriada à previsão da qualidade

do ar. Esses autores aplicaram a metodologia de Box-Jenkins para prever a qualidade do ar na cidade de Atenas, na Grécia.

Além de aplicações nas áreas citadas, Lim & McAleer (2001) mencionam, com base em diferentes métodos de previsão de séries temporais, que a literatura na área de previsão de demanda turística é numerosa. Os referidos autores aplicaram a metodologia de Box-Jenkins e modelaram a demanda de turistas que chegam à Austrália provenientes de três países da Ásia.

Neste artigo, a metodologia de Box-Jenkins é aplicada aos dados históricos de uma empresa que atua no setor de serviços, fornecendo assistência técnica para computadores pessoais. A principal contribuição deste artigo pode ser resumida em três aspectos: 1. apresentação passo a passo do emprego da metodologia; 2. uso sequencial de gráficos e testes estatísticos para subsidiar a construção e a escolha dos modelos mais adequados; e 3. ênfase na modelagem orientada a cada segmento de cliente, uma vez que eles são influenciados por diferentes fatores e, assim, apresentam comportamento distinto no que se refere à tendência e à sazonalidade.

## 2. Modelos Box-Jenkins

Conforme Morretin & Tolo (1987), uma série temporal é qualquer conjunto de observações ordenadas no tempo. Séries temporais são compostas por quatro elementos:

1. *Tendência*: verifica o sentido de deslocamento da série ao longo de vários anos.
2. *Ciclo*: movimento ondulatório que ao longo de vários anos tende a ser periódico.
3. *Sazonalidade*: movimento ondulatório de curta duração, em geral, inferior a um ano; associada, na maioria dos casos, a mudanças climáticas.
4. *Ruído aleatório ou erro*: compreende a variabilidade intrínseca aos dados e não pode ser modelado.

Como uma série temporal tem os dados coletados seqüencialmente ao longo do tempo, espera-se que ela apresente correlação seriada no tempo. Os modelos de Box-Jenkins, genericamente conhecidos por ARIMA (*Auto Regressive Integrated Moving Averages*) e na literatura em português por Auto-regressivos Integrados de Médias Móveis, são modelos matemáticos que visam captar o comportamento da correlação seriada ou autocorrelação entre os valores da série temporal, e com base nesse comportamento realizar previsões futuras. Se essa estrutura de correlação for bem modelada, fornecerá boas previsões.

Segundo Fava (2000), os modelos ARIMA resultam da combinação de três componentes denominados “filtros”: o componente auto-regressivo (AR), o filtro de integração (I) e o componente de médias móveis (MA). Uma série pode ser modelada pelos três filtros ou apenas um subconjunto deles, resultando em vários modelos abordados a seguir.

## 2.1 Modelos estacionários

Modelos estacionários são aqueles que assumem que o processo está em “equilíbrio”. Um processo é considerado fracamente estacionário se sua média e variância se mantêm constantes ao longo do tempo e a função de autocovariância depende apenas da defasagem entre os instantes de tempo. Um processo é fortemente estacionário se todos os momentos conjuntos são invariantes a translações no tempo.

### *Modelo auto-regressivo (AR)*

Em um modelo auto-regressivo, a série de dados históricos  $Z_t$  é descrita por seus valores passados regredidos e pelo ruído aleatório  $\varepsilon_t$ . Assim, um modelo AR(p) é dado por:

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + \varepsilon_t \quad (2.1)$$

em que:

$$\tilde{Z}_t = Z_t - \mu ;$$

$\phi_i$  é o parâmetro que descreve como  $\tilde{Z}_t$  se relaciona com o valor  $\tilde{Z}_{t-i}$  para  $i = 1, 2, \dots, p$ .

O modelo AR(p) dado pela equação 2.1 pode ser reescrito conforme apresentado na equação 2.2, utilizando o operador de defasagem L (aplicando o operador de defasagem em  $Z_{t-3}$  temos  $L^3 Z_t$ ).

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) \tilde{Z}_t = \varepsilon_t \quad (2.2)$$

O modelo auto-regressivo de ordem 1 ou AR(1) é a versão mais simples dessa classe de modelos. Sua apresentação algébrica é dada pela equação 2.3:

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \varepsilon_t \quad (2.3)$$

Para o modelo ser estacionário é necessário que  $|\phi_1| < 1$  (condição de estacionariedade) e que as autocovariâncias ( $\gamma_k$ ) sejam independentes. No caso do modelo AR(1), as autocovariâncias são dadas por:

$$\gamma_k = \phi_1^k \gamma_0 \quad (2.4)$$

e as autocorrelações  $\rho_k$  são dadas pela equação 2.5:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1^k \quad k = 0, 1, 2, \dots \quad (2.5)$$

A função de autocorrelação decai exponencialmente quando  $\phi_1$  é positivo; quando  $\phi_1$  é negativo, a função de autocorrelação também decai exponencialmente, mas apresenta alternância de sinais positivos e negativos.

### *Modelos de médias móveis (MA)*

Em um modelo de médias móveis (do inglês *moving average*), a série  $Z_t$  resulta da combinação dos ruídos brancos  $\varepsilon$  do período atual com aqueles ocorridos em períodos anteriores. Assim, um modelo de médias móveis de ordem q ou MA(q) é dado por:

$$\tilde{Z}_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (2.6)$$

em que:

$$\tilde{Z}_t = Z_t - \mu;$$

$\theta_i$  é o parâmetro que descreve como  $Z_t$  se relaciona com o valor  $\varepsilon_{t-i}$  para  $i = 1, 2, \dots, q$ .

O modelo MA(q) dado pela equação 2.6 pode ser reescrito, como em 2.7, utilizando o operador de defasagem L.

$$\begin{aligned} (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) \varepsilon_t &= \\ &= \theta(L) \varepsilon_t = \tilde{Z}_t \end{aligned} \quad (2.7)$$

O modelo MA(1) é a versão mais simples dessa classe de modelos. Sua apresentação algébrica é dada pela equação 2.8:

$$\tilde{Z}_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} \quad (2.8)$$

As autocorrelações  $\rho_k$ , que nada mais são do que as autocovariâncias divididas pela variância, são dadas por:

$$\begin{aligned} \rho_1 &= \frac{\gamma_1}{\gamma_0} = \frac{-\theta_1 \sigma_\varepsilon^2}{(1 + \theta_1^2) \sigma_\varepsilon^2} = \frac{-\theta_1}{(1 + \theta_1^2)} \text{ e} \\ \rho_k &= 0 \quad k > 1 \end{aligned} \quad (2.9)$$

A função de autocorrelação do modelo MA(1) apresenta apenas a primeira autocorrelação não-nula e as demais iguais a zero. A primeira autocorrelação será positiva se  $\theta_1$  for menor que zero e negativa se  $\theta_1$  for maior que zero.

Segundo Nelson (1973), uma propriedade importante da MA(1), proveniente da função de autocorrelação, é que sua “memória” é de somente um período. Uma dada observação, por exemplo  $Z_{53}$ , está correlacionada apenas a seu antecessor  $Z_{52}$  e a seu sucessor  $Z_{54}$ , mas não a qualquer outro membro da série.

Para obter a condição de invertibilidade, isto é, para transformar um modelo MA(1) em um modelo AR( $\infty$ ), é preciso impor a restrição de que  $|\theta_1| < 1$ . Segundo Abraham & Ledolter (1983), para que um processo MA(q) se torne inversível, é necessário que as raízes da equação:

$$\theta(L) = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) = 0$$

sejam maiores que um.

### Modelos auto-regressivos de médias móveis (ARMA)

Em alguns casos, pode ser necessário utilizar um grande número de parâmetros em modelos puramente AR ou puramente MA. Nesses casos, é vantajoso misturar os componentes de um modelo AR como os componentes de um modelo MA, gerando, assim, um modelo ARMA. O modelo ARMA(p,q) exigirá um número menor de termos e pode ser expresso conforme a equação 2.10:

$$\begin{aligned} \tilde{Z}_t &= \phi_1 \tilde{Z}_{t-1} + \dots + \phi_p \tilde{Z}_{t-p} + \varepsilon_t - \\ &- \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \end{aligned} \quad (2.10)$$

O modelo ARMA mais simples é o ARMA(1,1), dado pela equação 2.11:

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (2.11)$$

A função de autocorrelação do modelo ARMA(1,1) é dada por:

$$\begin{aligned} \rho_1 &= \frac{(1 - \phi_1 \theta_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 + 2\phi_1 \theta_1} \\ \rho_k &= \phi_1 \rho_{k-1} \text{ para } k > 1 \end{aligned} \quad (2.12)$$

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

## 2.2 Modelos não-estacionários

Quando uma série temporal apresenta média e variância dependentes do tempo, é porque ela não é estacionária. A não-estacionariedade de uma série implica que: a) há inclinação nos dados e eles não permanecem ao redor de uma linha horizontal ao longo do tempo e/ou b) a variação dos dados não permanece essencialmente constante sobre o tempo, isto é, as flutuações

aumentam ou diminuem com o passar do tempo, indicando que a variância está se alterando.

Para detectar a não-estacionariedade de uma série, o comportamento temporal pode ser analisado graficamente, buscando padrões (a) e (b) ou, então, aplicando os testes estatísticos de raiz unitária. O teste de raiz unitária mais usado é o de Dickey-Fuller. Outros detalhes sobre esse teste podem ser vistos em Makridakis *et al.* (1998), Enders (1995) e Hamilton (1994).

### *Modelos auto-regressivos integrados de médias móveis (ARIMA)*

Como a maioria dos procedimentos de análise estatística de séries temporais supõe que estas sejam estacionárias, será necessário transformá-las caso ainda não sejam. Segundo Morretin & Tolo (1987), a transformação mais comum consiste em tomar diferenças sucessivas da série original até obter uma série estacionária.

A primeira diferença de  $Z_t$  é definida por 2.13:

$$\Delta Z_t = Z_t - Z_{t-1} \quad (2.13)$$

a segunda é dada por 2.14:

$$\begin{aligned} D^2 Z_t &= D [D Z_t] = D[Z_t - Z_{t-1}] = \\ &= Z_t - 2Z_{t-1} - Z_{t-2} \end{aligned} \quad (2.14)$$

Em situações normais, ainda segundo os autores citados, será suficiente tomar uma ou duas diferenças para que a série se torne estacionária. O número  $d$  de diferenças necessárias para tornar a série estacionária é denominado ordem de integração. A inclusão do termo de ordem de integração permite que sejam utilizados os modelos ARIMA( $p,d,q$ ) dados pela equação 2.15:

$$\begin{aligned} w_t &= \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + \\ &+ \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \end{aligned} \quad (2.15)$$

em que:  $w_t = \Delta^d Z_t$ .

O modelo ARIMA( $p,d,q$ ) dado pela equação 2.15 pode ser reescrito, como em 2.16, utilizando o operador de defasagem  $L$ .

$$(1 - \phi_1 L - \dots - \phi_p L^p) w_t = (1 - \theta_1 L - \dots - \theta_q L^q) \varepsilon_t$$

com  $w_t = (1 - L)^d Z_t$

$$\phi(L) (1 - L)^d Z_t = \theta(L) \varepsilon_t \quad (2.16)$$

## 2.3 Modelos sazonais

Os modelos ARIMA exploram a autocorrelação entre os valores da série em instantes sucessivos, mas quando os dados são observados em períodos inferiores a um ano, a série também pode apresentar autocorrelação para uma estação de sazonalidade  $s$ . Os modelos que contemplam as séries que apresentam autocorrelação sazonal são conhecidos como SARIMA.

Os modelos SARIMA contêm uma parte não sazonal, com parâmetros ( $p,d,q$ ), e uma sazonal, com parâmetros ( $P,D,Q$ ) <sub>$s$</sub> . O modelo mais geral é dado pela equação 2.17:

$$\begin{aligned} (1 - \phi_1 L - \dots - \phi_p L^p) (1 - \Phi_1 L^s - \dots - \Phi_P L^{Ps}) \\ (1 - L)^d (1 - L^s)^D Z_t = (1 - \theta_1 L - \dots - \theta_q L^q) \\ (1 - \Theta_1 L^s - \dots - \Theta_Q L^{Qs}) \varepsilon_t \end{aligned} \quad (2.17)$$

em que:

$(1 - \phi_1 L - \dots - \phi_p L^p)$  é a parte auto-regressiva não-sazonal de ordem  $p$ ;

$(1 - \Phi_1 L^s - \dots - \Phi_P L^{Ps})$  é a parte auto-regressiva sazonal de ordem  $P$  e estação sazonal  $s$ ;

$(1 - L)^d$  é parte de integração não-sazonal de ordem  $d$ ;

$(1 - L^s)^D$  é parte de integração sazonal de ordem  $D$  e estação sazonal  $s$ ;

$(1 - \theta_1 L - \dots - \theta_q L^q)$  é a parte não-sazonal de médias móveis de ordem  $q$ ;

$(1 - \Theta_1 L^s - \dots - \Theta_Q L^{Qs})$  é a parte sazonal de médias móveis de ordem  $Q$  e estação sazonal  $s$ .

## 2.4 Etapas da metodologia Box-Jenkins

Segundo Morretin & Tolo (1987), a construção dos modelos Box-Jenkins é baseada em um ciclo iterativo, no qual a escolha do modelo é feita com base nos próprios dados. Segundo Box & Jenkins (1976), são três as etapas para construção do modelo:

1. *Identificação*: consiste em descobrir qual dentre as várias versões dos modelos de Box-Jenkins, sejam eles sazonais ou não, descreve o comportamento da série. A identificação do modelo a ser estimado ocorre pelo comportamento das funções de autocorrelações (ACF) e das funções de autocorrelações parciais (PACF). Outros detalhes referentes à obtenção dessas funções e a quais comportamentos representam os modelos anteriormente abordados podem ser pesquisados em Makridakis *et al.* (1998).
2. *Estimação*: consiste em estimar os parâmetros  $\phi$  e  $\Phi$  do componente auto-regressivo, os parâmetros  $\theta$  e  $\Theta$  do componente de médias móveis e a variância de  $\epsilon_t$ .
3. *Verificação*: consiste em avaliar se o modelo estimado é adequado para descrever o comportamento dos dados.

Caso o modelo não seja adequado, o ciclo é repetido, voltando-se à fase de identificação. Um procedimento muito utilizado é identificar não só um único modelo, mas alguns modelos que serão então estimados e verificados. Quando se obtém um modelo satisfatório, passa-se para a última etapa da metodologia de Box-Jenkins, que constitui o objetivo principal da metodologia: realizar previsões. O detalhamento dessas etapas se dará com a aplicação no estudo de caso.

## 3. Estudo de caso

Com o objetivo de realizar previsões de demanda, inicialmente, foram reunidos dados históricos de agosto de 1996 a julho de 2001 de

uma empresa prestadora de serviços de assistência técnica em equipamentos de informática. Por motivos estratégicos, os dados mais atuais não foram disponibilizados para publicação. No entanto, isso não compromete a apresentação deste artigo, cujo propósito é esclarecer a utilização da metodologia e discutir a importância da segmentação de clientes. Trata-se de uma empresa de pequeno porte que atualmente conta com duas lojas de atendimento aos clientes na capital gaúcha, com aproximadamente 40 funcionários. Essa empresa tem três tipos diferenciados de clientes: os clientes que assinam **contratos** de manutenção, em geral pessoas jurídicas; os clientes que trazem o equipamento para manutenção no período de **garantia**; e os clientes **avulsos**, em geral pessoas físicas. A empresa realiza serviços de manutenção de hardware e/ou software e, raramente, vende peças sem realizar algum tipo de serviço.

A demanda da empresa não pode ser estabelecida sem uma análise segmentada dos tipos de clientes. Cada tipo de cliente atendido pela empresa tem suas peculiaridades, por isso, a previsão de demanda total deve ser a soma da previsão da demanda individual dos tipos de clientes. Assim, para realizar a previsão de demanda da empresa foi empregada a metodologia de Box-Jenkins para o número de atendimentos realizados em cada um dos três tipos de clientes.

### 3.1 Metodologia Box-Jenkins para clientes tipo contrato

#### 3.1.1 Etapa de identificação

Para identificar os modelos apropriados, inicialmente deve ser analisado o gráfico de tempo da série em estudo. A análise desse gráfico pode indicar a presença de tendência ou alteração na variância, o que revelaria se a série é ou não estacionária. Quanto ao número de atendimentos de clientes tipo contrato, encontrado na Figura 1, observa-se fraca tendência decrescente. Portanto, como a série apresenta inclinação muito pequena, ela será considerada estacionária. Essa

suposição coincide com a expectativa da equipe técnica da empresa, que não acredita que possa ocorrer diminuição do número de contratos nos próximos meses.

O próximo passo é analisar as funções de autocorrelações (ACF) e de autocorrelações parciais (PACF). O comportamento dessas funções indica qual o modelo a ser utilizado, bem como auxilia no uso dos testes de raízes unitárias para confirmar a estacionariedade.

A Figura 2a apresenta a ACF do número de atendimentos de clientes tipo contrato. Pode-se observar que a ACF decai exponencialmente rápido, chegando a zero nas primeiras defasagens, indicação de um modelo AR(1).

A PACF apresenta as autocorrelações parciais, as quais medem a correlação entre o valor de  $Z_t$  e  $Z_{t-k}$ , descontadas a influência dos valores de  $Z_{t-1}$  até  $Z_{t-k+1}$ , e são obtidas pelas equações de Yule-Walker (Figura 2b). Observa-se que apenas a autocorrelação de defasagem 1 é significativa, reforçando a indicação de um modelo AR(1). Como as ACF e PACF fornecem a indicação de um modelo auto-regressivo de ordem um, testou-se a estacionariedade da série por intermédio do teste Dickey-Fuller, o mais conhecido teste de

raízes unitárias. Para tal, utilizou-se o pacote econométrico EViews (versão 3.0), em que a hipótese de não-estacionariedade foi rejeitada em nível de 5% de significância. Visto que a série pode ser considerada estacionária, a ordem de integração é zero, isto é,  $d = 0$ .

### 3.1.2 Etapa de estimativa

Uma vez indicados os valores de  $p$ ,  $d$  e  $q$ , passa-se para a estimativa dos parâmetros do modelo proposto, da variância do resíduo  $\sigma_\varepsilon^2$  e da constante do modelo que, neste caso, é o nível médio do processo.

Para estimar os parâmetros do modelo foi usado o pacote estatístico SPSS (versão 8.0), optando-se pelo método de *forecasting* de mínimos quadrados incondicional. As estimativas para o modelo em estudo foram:  $\phi_1 = 0,59$  e  $\mu = 95,59$ .

Reescrevendo o modelo dado pela equação 2.3 e substituindo os valores dos coeficientes, temos:

$$\begin{aligned}\tilde{Z}_t &= 0,59\tilde{Z}_{t-1} + \varepsilon_t \\ Z_t - 95,59 &= 0,59(Z_{t-1} - 95,59) + \varepsilon_t \\ Z_t &= 0,59Z_{t-1} + 39,20 + \varepsilon_t\end{aligned}\quad (3.1)$$

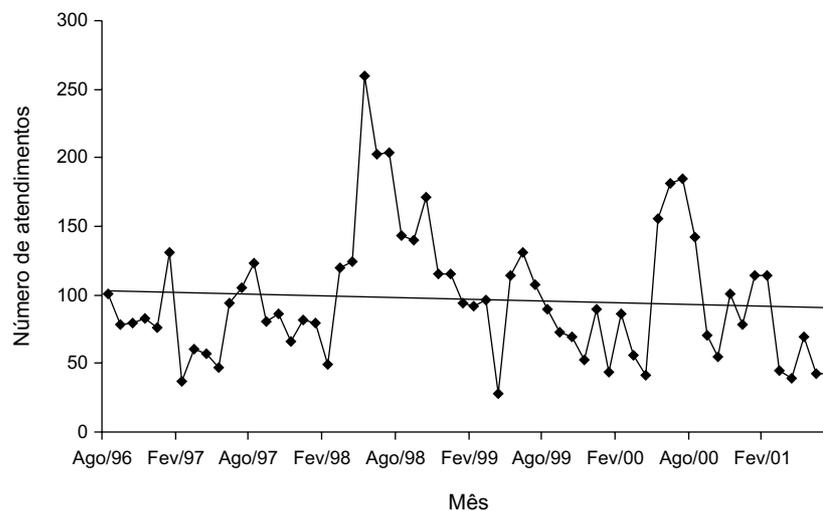
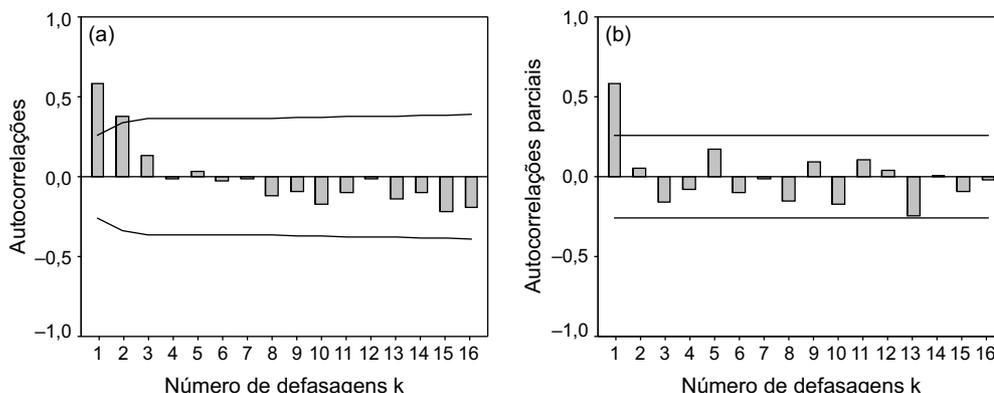


Figura 1 – Número de atendimentos de clientes tipo contrato e sua tendência.



**Figura 2 – a) Função de autocorrelação e b) função de autocorrelação parcial para o número de atendimentos de clientes tipo contrato.**

### 3.1.3 Etapa de verificação

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

#### *Avaliação por meio da análise de resíduos*

Os resíduos do modelo estimado  $\hat{\epsilon}_t$  são estimativas do ruído branco, sendo assim, devem apresentar esse comportamento se o modelo estiver adequadamente especificado, isto é, suas autocorrelações devem ser não-significantes. Para testar essa suposição, verificou-se o comportamento da ACF para os resíduos estimados. A análise da Figura 3 revela que os resíduos apresentam comportamento aleatório, ou seja, não revelam padrão específico. Assim, a análise dos resíduos confirma a adequação do modelo.

O uso do teste de Ljung-Box  $Q^*$  reforça essa afirmativa. Esse teste compara o valor da estatística de teste com os valores tabelados da distribuição Qui-quadrado com  $k-p-q$  graus de liberdade. Aplicando o teste para  $k = 15$  sobre os erros estimados pela aplicação do modelo AR(1), obtém-se uma

estatística de teste  $Q^* = 21,07$ , que não é significativa quando comparada ao valor da distribuição Qui-quadrado com  $k - p = 15 - 1 = 14$  graus de liberdade. Assim, pela análise do teste de Ljung-Box, os resíduos podem ser considerados ruído branco.

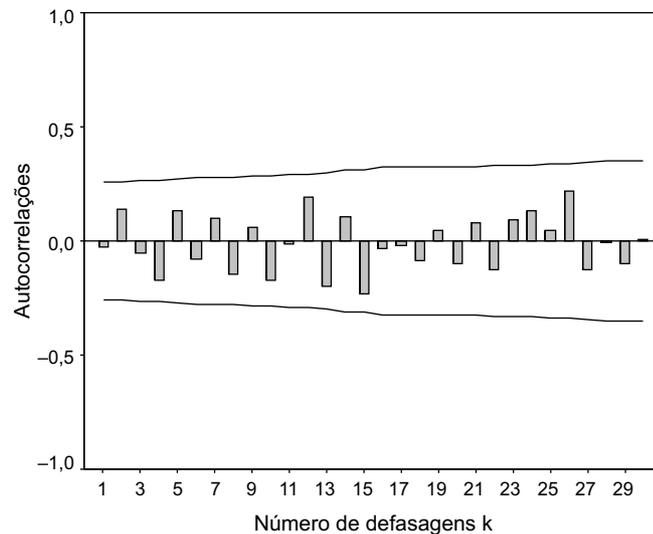
#### *Avaliação da ordem do modelo*

O objetivo é verificar se o modelo é parcimonioso, isto é, se não tem parâmetros em excesso. Essa verificação é realizada com base no erro-padrão dos coeficientes. Se o valor do coeficiente estimado for pequeno em relação a seu erro-padrão, conclui-se que ele não é significativo, ou seja, não há evidências estatísticas para suportar a inclusão do coeficiente no modelo.

Além disso, a análise do desvio-padrão residual é outro indicador para verificar se a ordem do modelo está adequada. Quanto menor o erro-padrão, melhores serão as previsões.

Os critérios de AIC (*Akaike Information Criteria*) e SBC (*Schwartz Bayesian Criteria*), utilizados para comparação de modelos e que levam em conta a variância do erro, o tamanho da amostra  $T$  e os valores de  $p$ ,  $q$ ,  $P$  e  $Q$ , também são utilizados. Makridakis *et al.* (1998) indicam que alguns programas computacionais utilizam para o cálculo aproximado de AIC a expressão:

$$T(1 + \ln(2p)) + T \ln s^2 + 2(p + q + P + Q)$$



**Figura 3 – Função de autocorrelação para os resíduos do modelo AR(1).**

A Tabela 1 apresenta uma comparação do modelo escolhido com outros dois modelos propostos para o número de atendimentos de clientes tipo contrato. Para o modelo AR(2) observa-se que o coeficiente  $\phi_2$  não é significativo (*p-value* igual a 0,661), logo, a variável no tempo  $t - 2$  não deve ser incluída no modelo. Para confirmar a estacionariedade da série, tomou-se uma diferença e foi estimado o modelo ARIMA(1,1,0) que apresenta desvio-padrão residual maior que o modelo indicado. Essas comparações levaram à escolha do modelo AR(1), apresentado na equação 3.1.

### 3.1.4 Etapa de previsão

Segundo Box & Jenkins (1976), para um modelo auto-regressivo de primeira ordem, o predictor  $h$  passo a frente é dado por:

$$\hat{Z}_t(h) = \phi_1^h Z_t + \mu - \mu \phi_1^h \quad (3.2)$$

em que:  $\hat{Z}_t(h)$  é a previsão  $h$  passos a frente com base no tempo  $t$ ;  $\phi_1$  é o coeficiente da variável em estudo defasada um período; e  $\mu$ , a média do processo.

**Tabela 1 – Critérios de comparação para verificação do melhor modelo.**

Modelo	<i>p-value</i> dos coeficientes	Critério AIC	Critério SBC	Desvio-padrão residual	Variância residual
AR(1)	$\phi_1$ : 0,000 nível do processo: 0,000	610,12 (T = 60)	614,31 (T = 60)	38,29	1.466,46
AR(2)	$\phi_1$ : 0,000 $\phi_2$ : 0,661 nível do processo: 0,000	611,95 (T = 60)	618,23 (T = 60)	38,56	1.486,84
ARIMA(1,1,0)	$\phi_1$ : 0,056 nível do processo: 0,832	609,47 (T = 59)	613,63 (T = 59)	41,62	1.732,37

Fonte: Análise no pacote SPSS, versão 8.0.

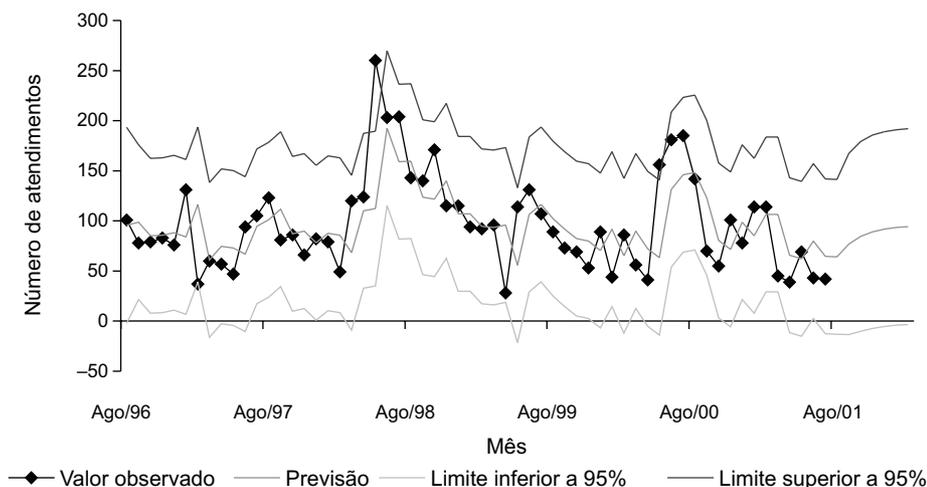
Para o estudo do número de atendimentos de clientes tipo contrato, a previsão com base no período de julho de 2001 que corresponde a  $t = 60$ , é dada pela equação 3.3:

$$\hat{Z}_{60}(h) = 0,59^h 42 + 95,59 - 95,59 * 0,59^h \quad (3.3)$$

A Tabela 2 apresenta as previsões para a demanda do número de atendimentos de clientes tipo contrato, de agosto de 2001 a fevereiro de 2002, obtidas com base no período de julho de 2001. O gráfico com as previsões e o intervalo de 95% de confiança para as previsões encontra-se na Figura 4.

**Tabela 2 – Previsões da demanda do número de atendimentos de clientes tipo contrato.**

Mês	T	h	$\hat{Z}_t(h)$
Agosto/2001	61	1	$\hat{Z}_{60}(1) = 0,59^1 * 42 + 95,59 - 95,59 * 0,59^1 = 63,97 = 64$
Setembro/2001	62	2	$\hat{Z}_{60}(2) = 0,59^2 * 42 + 95,59 - 95,59 * 0,59^2 = 76,94 = 77$
Outubro/2001	63	3	$\hat{Z}_{60}(3) = 0,59^3 * 42 + 95,59 - 95,59 * 0,59^3 = 84,59 = 85$
Novembro/2001	64	4	$\hat{Z}_{60}(4) = 0,59^4 * 42 + 95,59 - 95,59 * 0,59^4 = 89,10 = 89$
Dezembro/2001	65	5	$\hat{Z}_{60}(5) = 0,59^5 * 42 + 95,59 - 95,59 * 0,59^5 = 91,76 = 92$
Janeiro/2002	66	6	$\hat{Z}_{60}(6) = 0,59^6 * 42 + 95,59 - 95,59 * 0,59^6 = 93,33 = 94$
Fevereiro/2002	67	7	$\hat{Z}_{60}(7) = 0,59^7 * 42 + 95,59 - 95,59 * 0,59^7 = 94,26 = 95$



**Figura 4 – Previsão da demanda do número de atendimentos de clientes tipo contrato.**

## 3.2 Metodologia Box-Jenkins para clientes tipo garantia

### 3.2.1 Etapa de identificação

Para identificar os modelos apropriados, inicialmente, foi analisada a Figura 5, que apresenta o gráfico de tempo do número de atendimentos de clientes tipo garantia. A análise dessa figura indica a presença de tendência (crescimento), o que revela uma série não-estacionária. Além disso, a série apresenta um nível médio que deve ser considerado no modelo a ser proposto.

O próximo passo é analisar as funções de autocorrelações (ACF) e autocorrelações parciais (PACF) que podem auxiliar na verificação da estacionariedade e na definição do modelo mais apropriado.

A Figura 6a apresenta a ACF do número de atendimentos de clientes tipo garantia e revela um comportamento senoidal, típico do processo auto-regressivo.

Já a PACF mostrada na Figura 6b apresenta as três primeiras defasagens significativas, reforçando a indicação de um modelo auto-regressivo, no caso, um AR(3).

As ACF e PACF indicam um modelo auto-regressivo de ordem três, contudo, é preciso verificar a estacionariedade da série. A verificação foi realizada por meio do teste Dickey-Fuller.

Para tal, utilizou-se o pacote EViews (versão 3.0), em que a hipótese de não-estacionariedade não foi rejeitada em nível de 5% de significância. Assim, em virtude da evidência de não-estacionariedade da série, é necessário tomar a diferença. Utilizando a primeira diferença e reaplicando o teste Dickey-Fuller verifica-se que a série é estacionária. Assim, tem-se indicações de que a ordem de integração deve ser um, isto é,  $d = 1$ .

Para obter a estacionariedade da série é preciso trabalhar com a diferença, portanto, o modelo a ser indicado é um ARIMA(3,1,0).

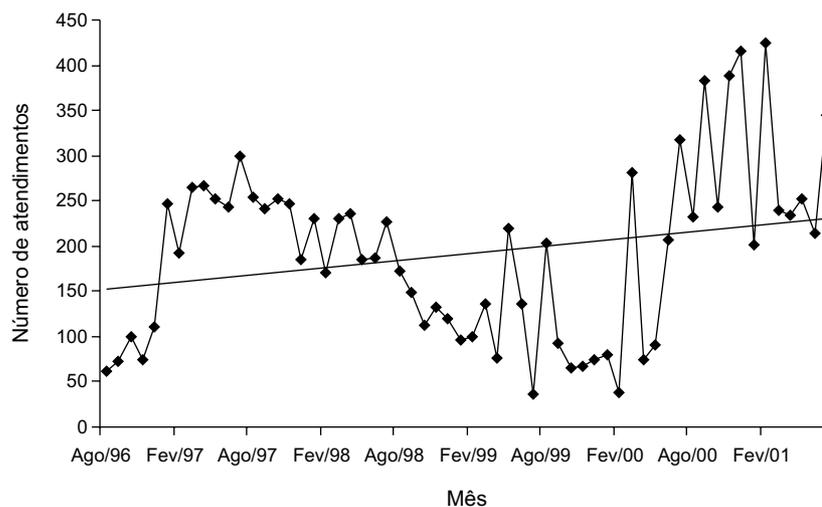
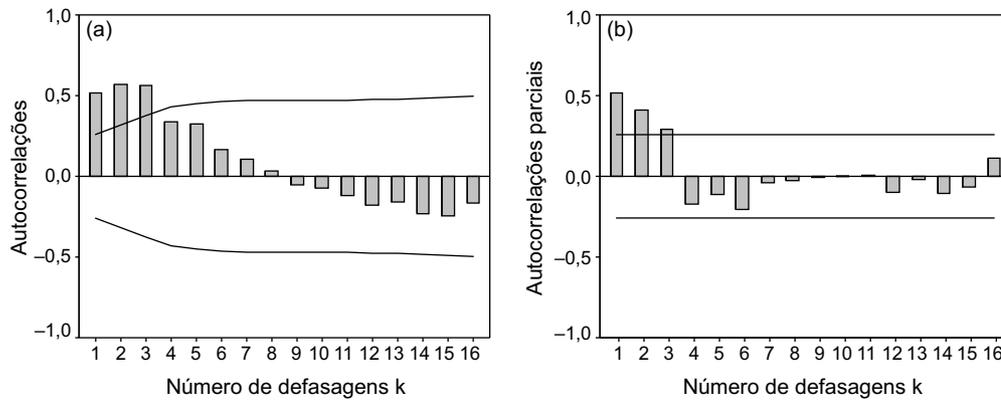


Figura 5 – Número de atendimentos de clientes tipo garantia e sua tendência.



**Figura 6 – a) Função de autocorrelação e b) função de autocorrelação parcial para o número de atendimentos de clientes tipo garantia.**

### 3.2.2 Etapa de estimativa

Uma vez indicados os valores de  $p$ ,  $d$  e  $q$ , passa-se para a estimativa dos parâmetros do modelo proposto e da variância do resíduo  $\sigma_{\varepsilon}^2$ . Para estimar os parâmetros do modelo foi utilizado o pacote estatístico SPSS (versão 8.0), optando-se pelo método de *forecasting* de mínimos quadrados incondicional.

Conforme será visto na próxima seção, o modelo ARIMA(3,1,0) indicado anteriormente não é o mais adequado. O modelo que melhor representa a série do número de atendimentos de clientes tipo garantia é o modelo ARIMA (2,1,0). As estimativas dos coeficientes para esse modelo foram:  $\phi_1 = -0,7865$ ;  $\phi_2 = -0,3795$ . Reescrevendo o modelo dado pela equação 2.16 (sem o termo constante), com  $p = 2$ ,  $d = 1$  e  $q = 0$ , e substituindo os valores dos coeficientes, tem-se:

$$(1 - \phi_1 L - \phi_2 L^2)(1 - L)Z_t = \varepsilon_t$$

$$Z_t = 0,2135Z_{t-1} + 0,407Z_{t-2} + 0,3795Z_{t-3} + \varepsilon_t \quad (3.4)$$

### 3.2.3 Etapa de verificação

#### *Avaliação por meio da análise de resíduos*

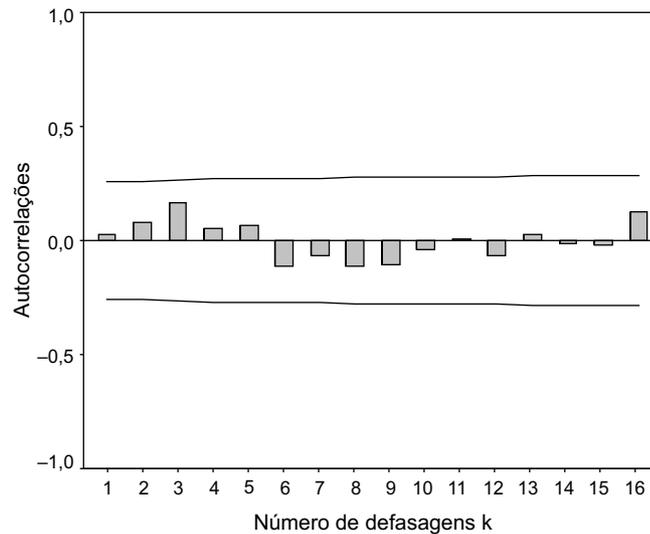
Se o modelo estiver adequadamente especificado, os resíduos do modelo estimado serão estimativas de ruído branco e, assim, os coeficientes de autocorrelação dos resíduos devem ser estatisticamente iguais a zero. A Figura 7 revela que os resíduos efetivamente apresentam comportamento aleatório (ruído branco), logo, o modelo é adequado no que se refere à análise dos resíduos. Essa conclusão é reforçada após a utilização do teste de Ljung-Box  $Q^*$ , que não rejeita, para qualquer defasagem  $k$ , a hipótese de erros não correlacionados. Assim, podemos considerar que os resíduos se comportam como ruído branco.

#### *Avaliação da ordem do modelo*

A Tabela 3 apresenta uma comparação do modelo proposto ARIMA(3,1,0) com outros três modelos para o número de atendimentos dos clientes tipo garantia. A análise do modelo ARIMA(3,1,0) revela que o coeficiente  $\phi_3$  não é significativo ( $p$ -value igual a 0,358). Como o

terceiro componente não é significativo, verificou-se o modelo ARIMA(2,1,0). Esse modelo, por sua vez, apresentou o menor desvio-padrão residual, assim como os menores valores dos critérios

AIC e SBC, quando comparado a modelos menos parcimoniosos que constam na tabela. Essas comparações levaram à escolha do modelo ARIMA(2,1,0) apresentado na equação 3.4.



**Figura 7 – Função de autocorrelação para os resíduos do modelo ARIMA(2,1,0).**

**Tabela 3 – Critérios de comparação para verificação do melhor modelo.**

Modelo	<i>p-value</i> dos coeficientes	Critério AIC	Critério SBC	Desvio-padrão residual	Variância residual
ARIMA(3,1,0)	$\phi_1$ : 0,000 $\phi_2$ : 0,092 $\phi_3$ : 0,358	673,38 (T = 59)	679,61 (T = 59)	70,54	4.976,19
ARIMA(2,1,0)	$\phi_1$ : 0,000 $\phi_2$ : 0,003	672,28 (T = 59)	676,43 (T = 59)	70,50	4.969,91
ARIMA(1,1,1)	$\phi_1$ : 0,063 $\theta_1$ : 0,068	676,28 (T = 59)	680,43 (T = 59)	73,03	5.333,84
ARIMA(2,1,1)	$\phi_1$ : 0,000 $\phi_2$ : 0,004 $\theta_1$ : 0,405	673,57 (T = 59)	679,80 (T = 59)	70,66	4.993,00

Fonte: Análise no pacote SPSS, versão 8.0.

### 3.2.4 Etapa de previsão

Para o modelo ARIMA(2,1,0) estimado, o valor  $(t + h)$  é dado por:

$$Z_{t+h} = 0,2135Z_{t+h-1} + 0,407Z_{t+h-2} + 0,3795Z_{t+h-3} + \varepsilon_{t+h} \quad (3.5)$$

Assim, o previsor  $h$  passos a frente para o estudo do número de atendimentos de clientes tipo contrato, é dado pelo seguinte conjunto de equações:

$$\begin{aligned} \hat{Z}_t(1) &= 0,2135Z_t + 0,407Z_{t-1} + 0,3795Z_{t-2} \\ \hat{Z}_t(2) &= 0,2135Z_{t+1} + 0,407Z_t + 0,3795Z_{t-1} = \\ &= 0,2135\hat{Z}_t(1) + 0,407Z_t + 0,3795Z_{t-1} \\ \hat{Z}_t(3) &= 0,2135Z_{t+2} + 0,407Z_{t+1} + 0,3795Z_t = \\ &= 0,2135\hat{Z}_t(2) + 0,407\hat{Z}_t(1) + 0,3795Z_t \\ \hat{Z}_t(h) &= 0,2135\hat{Z}_t(h-1) + 0,407\hat{Z}_t(h-2) + \\ &+ 0,3795\hat{Z}_t(h-3) \quad h \geq 4 \quad (3.6) \end{aligned}$$

A Tabela 4 apresenta as previsões para a demanda do número de atendimentos de clientes tipo garantia, de agosto de 2001 a fevereiro de 2002, com base no período de julho de 2001, que corresponde a  $t = 60$ . O gráfico com as previsões e o intervalo de 95% de confiança para as previsões encontra-se na Figura 8.

## 3.3 Metodologia Box-Jenkins para clientes tipo avulsos

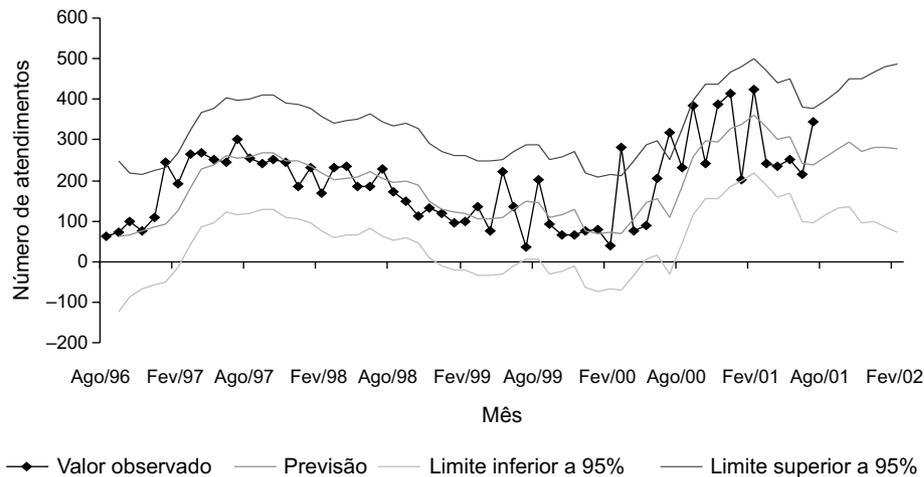
### 3.3.1 Etapa de identificação

A análise da Figura 9 auxilia a identificar os possíveis modelos a serem utilizados. Essa figura revela a presença de tendência crescente, o que indica que a série não é estacionária. Além disso, a série apresenta valor médio a ser considerado no modelo a ser proposto (constante).

A seguir, é realizada a análise das funções de autocorrelações (ACF) e autocorrelações parciais (PACF), que auxiliam na verificação da estacionariedade e na proposição de um modelo.

**Tabela 4 – Previsões da demanda do número de atendimentos de clientes tipo garantia.**

Mês	t	h	$\hat{Z}_t(h)$
Agosto/2001	61	1	$\hat{Z}_{60}(1) = 0,2135 * Z_{60} + 0,407 * Z_{59} + 0,3795 * Z_{58}$ $\hat{Z}_{60}(1) = 0,2135 * 344 + 0,407 * 215 + 0,3795 * 252 \cong 257$
Setembro/2001	62	2	$\hat{Z}_{60}(2) = 0,2135 * \hat{Z}_{60}(1) + 0,407 * Z_{60} + 0,3795 * Z_{59}$ $\hat{Z}_{60}(2) = 0,2135 * 256,58 + 0,407 * 344 + 0,3795 * 215 \cong 277$
Outubro/2001	63	3	$\hat{Z}_{60}(3) = 0,2135 * \hat{Z}_{60}(2) + 0,407 * \hat{Z}_{60}(1) + 0,3795 * Z_{60}$ $\hat{Z}_{60}(3) = 0,2135 * 276,38 + 0,407 * 256,58 + 0,3795 * 344 \cong 294$
Novembro/2001	64	4	$\hat{Z}_t(4) = 0,2135 * \hat{Z}_t(3) + 0,407 * \hat{Z}_t(2) + 0,3795 * \hat{Z}_t(1)$ $\hat{Z}_t(4) = 0,2135 * 293,98 + 0,407 * 276,38 + 0,3795 * 256,58 \cong 273$
Dezembro/2001	65	5	$\hat{Z}_{60}(5) = 0,2135 * \hat{Z}_{60}(4) + 0,407 * \hat{Z}_{60}(3) + 0,3795 * \hat{Z}_{60}(2)$ $\hat{Z}_{60}(5) = 0,2135 * 272,62 + 0,407 * 293,98 + 0,3795 * 276,38 \cong 283$
Janeiro/2002	66	6	$\hat{Z}_{60}(6) = 0,2135 * \hat{Z}_t(5) + 0,407 * \hat{Z}_t(4) + 0,3795 * \hat{Z}_t(3)$ $\hat{Z}_{60}(6) = 0,2135 * 282,74 + 0,407 * 272,62 + 0,3795 * 293,98 \cong 283$
Fevereiro/2002	67	7	$\hat{Z}_{60}(7) = 0,2135 * \hat{Z}_t(6) + 0,407 * \hat{Z}_t(5) + 0,3795 * \hat{Z}_t(4)$ $\hat{Z}_{60}(7) = 0,2135 * 282,89 + 0,407 * 282,74 + 0,3795 * 272,62 \cong 279$



**Figura 8 – Previsões da demanda do número de atendimentos de clientes tipo garantia.**

A Figura 10a, contendo a ACF do número de atendimentos de clientes tipo avulsos, revela que as autocorrelações apresentam decaimento exponencial, típico do processo auto-regressivo.

A Figura 10b contém a PACF, a qual apresenta apenas a autocorrelação parcial de ordem um significativa. Assim, há indicação de que a ordem do modelo auto-regressivo é um, reforçando a indicação de um modelo auto-regressivo, no caso, um modelo AR(1).

É preciso verificar a estacionariedade da série, realizada por meio do teste Dickey-Fuller, utilizando o pacote econométrico EViews (versão 3.0). A hipótese de não-estacionariedade não foi rejeitada em nível de 5% de significância. Isso indica que, em virtude da evidência de não-estacionariedade da série, é necessário utilizar a diferença. Utilizando a primeira diferença e reaplicando o teste Dickey-Fuller verifica-se que a série é estacionária em nível de 5%. Assim, tem-se a indicação de que a ordem de integração deva ser um, isto é,  $d = 1$ .

Com base nas análises das ACF, PACF e dos testes para verificar a estacionariedade da série, o modelo indicado é um ARIMA(1,1,0).

### 3.3.2 Etapa de estimativa

Para estimar os parâmetros do modelo foi utilizado o pacote estatístico SPSS (versão 8.0).

Além dos parâmetros do modelo, foi estimada a variância dos resíduos  $\sigma_{\epsilon}^2$ .

O modelo ARIMA(1,1,0) indicado não foi o mais adequado, como será visto na próxima seção. O modelo que melhor representa a série do número de atendimentos para clientes tipo avulsos é o modelo SARIMA (0,1,0) (0,1,1)<sub>12</sub>, pois a influência da sazonalidade teve de ser considerada. A estimativa do coeficiente de média móvel sazonal obtida para esse modelo foi:  $\Theta_1 = 0,6606$ . Reescrevendo o modelo dado pela equação 2.17, com  $D = 1$  e  $Q = 1$ , e substituindo o valor do coeficiente, tem-se:

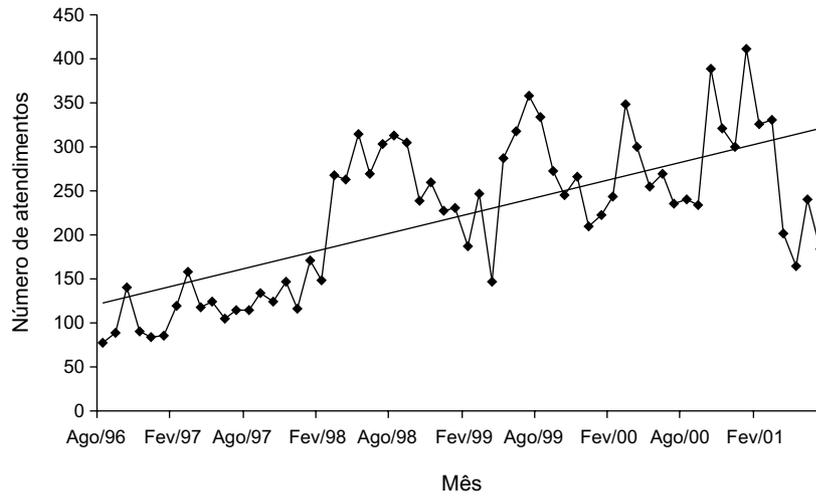
$$(1-L)^1(1-L^{12})^1 Z_t = (1-\Theta_1 L^{12}) \epsilon_t$$

$$Z_t = Z_{t-1} + Z_{t-12} - Z_{t-13} + \epsilon_t - 0,6606 \epsilon_{t-12} \quad (3.7)$$

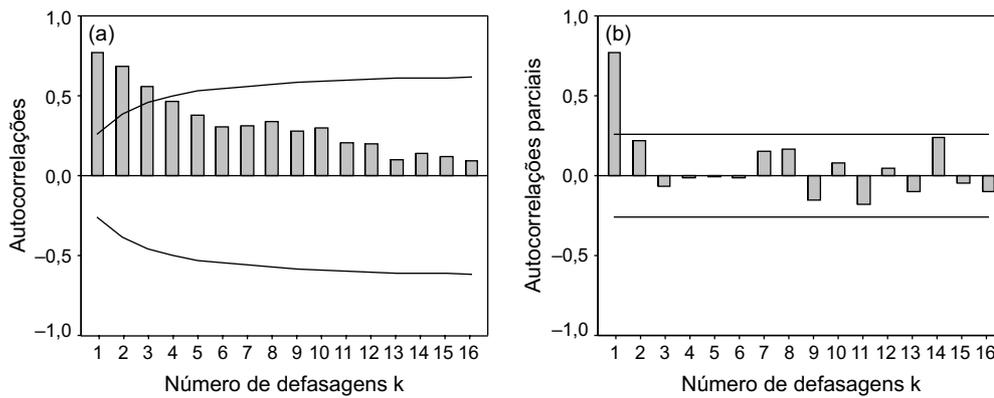
### 3.3.3 Etapa de verificação

*Avaliação por meio da análise de resíduos*

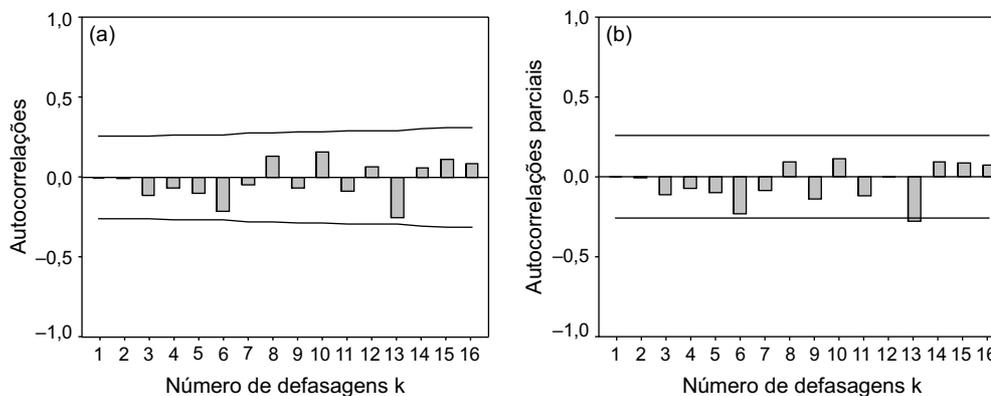
O modelo proposto pela análise da ACF, PACF e estacionariedade foi o ARIMA(1,1,0); contudo, como pode ser verificado na Figura 11, os resíduos não se comportam como ruído branco, pois a PACF revela a defasagem 13 significativa, o que indica a presença de sazonalidade, já referida na Figura 10a.



**Figura 9 – Número de atendimentos de clientes tipo avulsos e sua tendência.**



**Figura 10 – a) Função de autocorrelação e b) função de autocorrelação parcial para o número de atendimentos de clientes tipo avulsos.**



**Figura 11 – a) Função de autocorrelação e b) função de autocorrelação parcial para os resíduos do modelo ARIMA(1,1,0).**

*Avaliação da ordem do modelo*

Uma vez que os resíduos do modelo proposto indicaram a presença de sazonalidade, foram analisadas as ACF e PACF da primeira diferença em busca de outro modelo. A Figura 12 apresenta as ACF e PACF da primeira diferença, que indicam um modelo auto-regressivo sazonal. A indicação de termos auto-regressivos sazonais vem da significância das autocorrelações parciais de defasagens 1 e 13. Assim, foi proposto um modelo SARIMA(0,1,0)(1,0,0)<sub>12</sub>. Contudo, esse modelo apresentou coeficiente não-significativo para o termo auto-regressivo sazonal. Visando remover a influência do componente de sazonalidade, utilizou-se também a diferença sazonal de tamanho  $s = 12$ . Então, foi verificado o comportamento do modelo SARIMA(0,1,0)(1,1,0)<sub>12</sub>, que também apresentou coeficiente não-significativo, parecendo não ser o modelo mais adequado para este estudo.

A Tabela 5 apresenta alguns dos modelos analisados e os critérios que levaram à escolha do modelo SARIMA(0,1,0)(0,1,1)<sub>12</sub>. Dentre os modelos sazonais abordados, todos apresentam coeficientes não-significativos.

Os demais modelos não relacionados aqui ou apresentavam coeficiente não-significativo ou tinham desvio-padrão residual maior que o modelo SARIMA(0,1,0)(0,1,1)<sub>12</sub> escolhido. As autocorrelações residuais do modelo SARIMA

(0,1,0)(0,1,1)<sub>12</sub> são estatisticamente não-significativas. Isso pode ser afirmado com base na conclusão obtida pelo teste de Ljung-Box, que não rejeita, para qualquer defasagem  $k$ , a hipótese de erros não correlacionados. Assim, pode-se considerar, com base no teste Ljung-Box, que os resíduos comportam-se como ruído branco, sendo este modelo adequado para realizar previsões.

**3.3.4 Etapa de previsão**

Para o modelo SARIMA(0,1,0)(0,1,1)<sub>12</sub> estimado, o valor  $(t + h)$  é dado por:

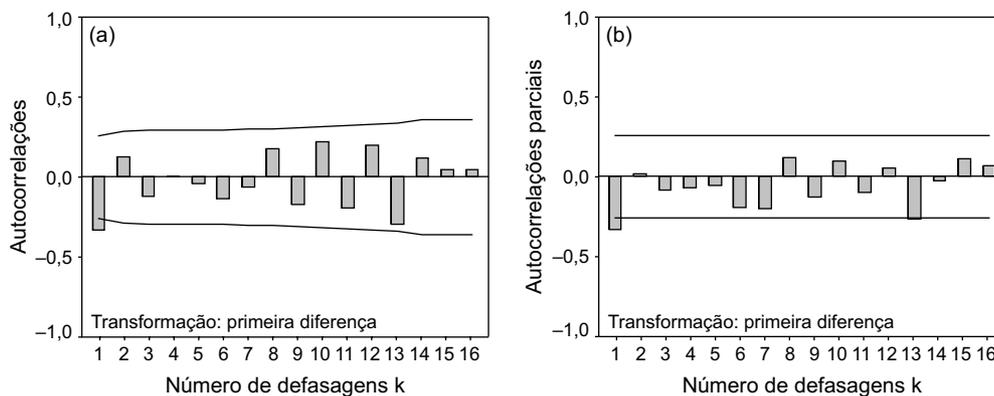
$$Z_{t+h} = Z_{t+h-1} + Z_{t+h-12} - Z_{t+h-13} + \epsilon_{t+h} - 0,6606 \cdot \epsilon_{t+h-12} \tag{3.8}$$

Assim, o previsor  $h$  passos a frente para o estudo do número de atendimentos de clientes tipo contrato é dado pela equação 3.9. O valor de  $\epsilon_{t+h}$ , que ainda não ocorreu, é estimado por sua média, ou seja, o valor zero:

$$\hat{Z}_t(1) = Z_{t+h-1} + Z_{t+h-12} - Z_{t+h-13} - 0,6606 \cdot \epsilon_{t+h-12} \tag{3.9}$$

$$\hat{Z}_t(h) = \hat{Z}_t(h-1) + Z_{t+h-12} - Z_{t+h-13} - 0,6606 \cdot \epsilon_{t+h-12} \tag{3.10}$$

sendo:  $2 \leq h \leq 12$ .



**Figura 12 – a) Função de autocorrelação e b) função de autocorrelação parcial para a primeira diferença da série do número de atendimentos de clientes tipo avulsos.**

**Tabela 5 – Critérios de comparação para verificar o melhor modelo.**

Modelo	<i>p-value</i> dos coeficientes	Critério AIC	Critério SBC	Desvio-padrão residual	Variância residual
ARIMA(1,1,0)	$\phi_1: 0,110$	636,61 (T = 59)	638,68 (T = 59)	52,81	2.788,45
SARIMA(0,1,0)(1,0,0) <sub>12</sub>	$\Phi_1: 0,100$	639,90 (T = 59)	641,99 (T = 59)	53,92	2.907,67
SARIMA(0,1,0)(1,1,0) <sub>12</sub>	$\Phi_1: 0,075$	525,29 (T = 47)	527,14 (T = 47)	63,08	3.975,11
SARIMA(1,1,0)(1,1,0) <sub>12</sub>	$\phi_1: 0,495$ $\Phi_1: 0,053$	526,83 (T = 47)	530,53 (T = 47)	63,84	3.999,18
SARIMA(0,1,0)(0,0,1) <sub>12</sub>	$\Theta_1: 0,167$	640,30 (T = 59)	642,38 (T = 59)	54,19	2.936,96
<b>SARIMA(0,1,0)(0,1,1)<sub>12</sub></b>	<b><math>\Theta_1: 0,056</math></b>	<b>523,12</b> (T = 47)	<b>524,97</b> (T = 47)	<b>58,03</b>	<b>3.367,17</b>
SARIMA(1,1,0)(1,0,0) <sub>12</sub>	$\phi_1: 0,034$ $\Phi_1: 0,487$	673,57 (T = 59)	679,80 (T = 59)	52,96	2.805,36

Fonte: Análise no pacote SPSS, versão 8.0.

A Tabela 6 apresenta as previsões para a demanda do número de atendimentos de clientes tipo avulsos, de agosto de 2001 a fevereiro de 2002. Essas previsões foram obtidas aplicando-se as equações 3.9 e 3.10 e tendo por base o período de julho de 2001, que corresponde a  $t = 60$ . O gráfico com as previsões e o intervalo de 95% de confiança para as previsões encontra-se na Figura 13.

### 3.4 Previsão da demanda futura da empresa

A previsão de demanda total da empresa, apresentada na Tabela 7, é a soma da previsão da demanda individual dos três tipos de clientes. As previsões apresentadas referem-se ao período de agosto de 2001 a fevereiro de 2002.

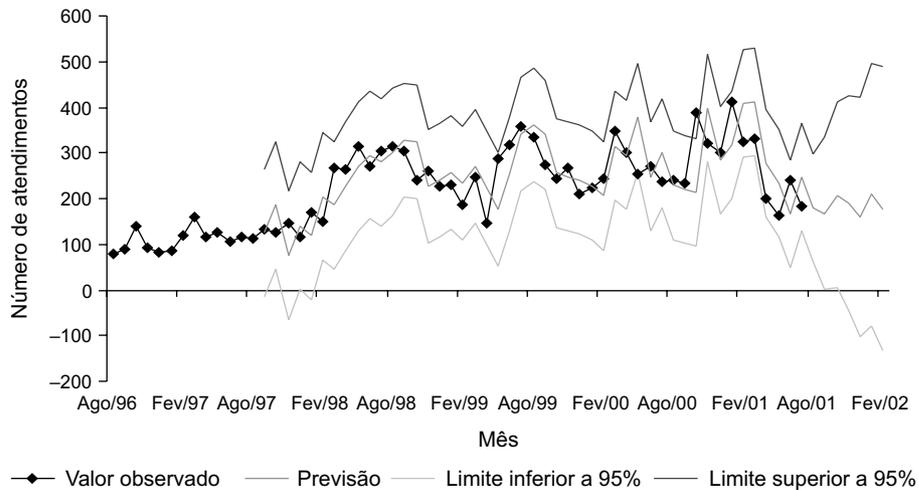
É importante observar que a previsão da demanda neste estudo de caso deve ser feita individualmente para cada segmento. A modelagem individual é essencial, na medida em que os dife-

rentes segmentos apresentam suas especificidades referentes à tendência, à sazonalidade e à ordem dos modelos. Vale mencionar que séries simples (contendo apenas tendência, por exemplo) podem ser modeladas com um número relativamente pequeno de observações, mas séries complexas (contendo termos de sazonalidade com múltiplos períodos, por exemplo) podem exigir 80 ou mais observações para obter boa modelagem. No caso em estudo, foram utilizadas 60 observações, pois representam a totalidade dos dados disponibilizados pela empresa.

Também vale mencionar que, na fase de verificação da metodologia Box-Jenkins, é comum reservar uma parte final dos dados disponíveis para verificar o desempenho do modelo, comparando as previsões com os valores observados. Neste estudo, isso não foi feito por duas razões: a amostra disponível (60 observações) era relativamente pequena e, por motivos estratégicos, a empresa não autorizou a divulgação de dados mais atuais.

**Tabela 6 – Previsões da demanda do número de atendimentos de clientes tipo avulsos.**

Mês	T	h	$\hat{Z}_t(h)$
Agosto/2001	61	1	$\hat{Z}_{60}(1) = Z_{60+1-1} + Z_{60+1-12} - Z_{60+1-13} - 0,6606 \cdot \epsilon_{60+1-12}$ $\hat{Z}_{60}(1) = 184 + 240 - 236 + 0,6606 * 11,12 \cong 181$
Setembro/2001	62	2	$\hat{Z}_{60}(2) = \hat{Z}_{60}(2-1) + Z_{60+2-12} - Z_{60+2-13} - 0,6606 \cdot \epsilon_{60+2-12}$ $\hat{Z}_{60}(2) = 180,65 + 234 - 240 - 0,6606 * 12,94 \cong 166$
Outubro/2001	63	3	$\hat{Z}_{60}(3) = \hat{Z}_{60}(3-1) + Z_{60+3-12} + Z_{60+3-13} - 0,6606 \cdot \epsilon_{60+3-12}$ $\hat{Z}_{60}(3) = 166,10 + 389 - 234 - 0,6606 * 175,87 \cong 205$
Novembro/2001	64	4	$\hat{Z}_{60}(4) = \hat{Z}_{60}(4-1) + Z_{60+4-12} + Z_{60+4-13} - 0,6606 \cdot \epsilon_{60+4-12}$ $\hat{Z}_t(4) = 204,92 + 321 - 389 - 0,6606 * (-77,40) \cong 188$
Dezembro/2001	65	5	$\hat{Z}_{60}(5) = \hat{Z}_{60}(5-1) + Z_{60+5-12} + Z_{60+5-13} - 0,6606 \cdot \epsilon_{60+5-12}$ $\hat{Z}_{60}(5) = 188 + 300 - 321 - 0,6606 * 15,87 \cong 157$
Janeiro/2002	66	6	$\hat{Z}_{60}(6) = \hat{Z}_{60}(6-1) + Z_{60+6-12} + Z_{60+6-13} - 0,6606 \cdot \epsilon_{60+6-12}$ $\hat{Z}_{60}(6) = 156,52 + 411 - 300 - 0,6606 * 94,42 \cong 206$
Fevereiro/2002	67	7	$\hat{Z}_{60}(7) = \hat{Z}_{60}(7-1) + Z_{60+7-12} + Z_{60+7-13} - 0,6606 \cdot \epsilon_{60+7-12}$ $\hat{Z}_{60}(7) = 205,15 + 326 - 411 - 0,6606 * (-82,72) \cong 175$



**Figura 13 – Previsões da demanda do número de atendimentos de clientes tipo avulsos.**

**Tabela 7 – Previsão da demanda total da empresa e por tipo de cliente.**

Mês	Previsão de demanda			Total
	Contrato	Garantia	Avulsos	
Agosto/2001	64	257	181	502
Setembro/2001	77	277	166	520
Outubro/2001	85	294	205	584
Novembro/2001	89	273	188	550
Dezembro/2001	92	283	157	532
Janeiro/2002	94	283	206	583
Fevereiro/2002	95	279	175	549

#### 4. Conclusões

Realizar previsão de demanda é uma atividade importante, pois pode revelar as tendências de mercado e contribuir no planejamento estratégico da empresa. As previsões de demanda também auxiliam na solução de problemas mais imediatos, como a definição do número de técnicos necessários para atender o cliente final.

Visando alcançar esses objetivos, este trabalho utilizou a metodologia de Box-Jenkins e discutiu sua aplicação por meio de um estudo de caso dirigido ao setor de assistência técnica em computadores pessoais. No estudo de caso desenvolvido, a demanda da empresa foi segmentada pelo tipo de clientes: contrato, garantia e avulsos. Verificou-se que cada um desses segmentos apresenta características próprias. Como cada cliente tem suas peculiaridades, o comportamento de cada um é distinto no que se refere à tendência e à sazonalidade, exigindo um modelo diferente em sua representação.

A aplicação da metodologia de Box-Jenkins requer do pesquisador atenção e prática para escolher o modelo que melhor define o comportamento temporal dos dados. Com a aplicação dessa metodologia, apoiada em testes estatísticos para subsidiar a modelagem, foi possível definir modelos apropriados para os três segmentos de clientes.

Para modelar o número de atendimento de clientes tipo contrato foi utilizado modelo auto-

regressivo de primeira ordem, AR(1), um modelo simples e amplamente estudado na literatura. Para modelar o número de atendimento de clientes tipo garantia foi necessário um modelo com mais parâmetros, contando com termos diferenciais e auto-regressivo de segunda ordem, ARIMA(2,1,0). Por fim, o número de atendimento de clientes tipo avulsos revelou um comportamento mais complexo, com sazonalidade mensal. Nesse caso, foi necessário utilizar um modelo contendo termos diferenciais e sazonais, SARIMA(0,1,0)(0,1,1)<sub>12</sub>.

A modelagem utilizada captou as diferenças existentes entre os segmentos de clientes. Essas diferenças podem ser justificadas tecnicamente. Para os clientes tipo contrato a influência temporal é reduzida, pois a demanda e a concorrência parecem crescer na mesma taxa. O resultado é que o número de contratos da empresa varia pouco ao longo do ano. Já os clientes tipo garantia revelam maior dependência temporal, pois as alterações no mercado e os acordos de assistência técnica se refletem gradativamente nos negócios da empresa, ou seja, as mudanças demoram algum tempo para surtirem efeito em campo. Os clientes tipo avulsos, por sua vez, apresentam sazonalidade anual em função dos hábitos de aquisição (presentes de natal e retorno às aulas, por exemplo), os quais estão intimamente relacionados aos períodos do ano.

Por fim, vale mencionar que o trabalho realizado foi de grande utilidade para a empresa, principalmente por dois motivos: a modelagem

permitiu maior entendimento do comportamento dos diversos clientes; e os resultados foram imediatamente assimilados pela empresa,

constituindo fonte adicional de informação no suporte às decisões referentes a investimentos e dimensionamento da equipe técnica.

### **Referências Bibliográficas**

- ABDEL-AAL, R. E.; AL-GARNI, Z. Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis. *Energy*, v. 22, n. 11, p. 1059-1069, nov. 1997.
- ABRAHAM, B.; LEDOLTER, J. *Statistical methods for forecasting*. New York: John Wiley & Sons, 1983.
- BOX, G. E. P.; JENKINS, G. M. *Time series analysis forecasting and control*. San Francisco: Holden-Day, 1976. Edição revisada.
- ENDERS, W. *Applied econometric time series*. New York: John Wiley & Sons, 1995.
- FAVA, V. L. *Manual de econometria*. In: VASCONCELOS, M. A. S.; ALVES, D. São Paulo: Editora Atlas, 2000.
- HAMILTON, J. D. *Time series analysis*. New Jersey: Princeton University Press, 1994.
- LIM, C.; McALEER, M. Time series forecasts of international travel demand for australia. *Tourism Management*. (Artigo aceito em julho de 2001, aguarda impressão.)
- MAKRIDAKIS, S.; WHEELWRIGHT, S.; HYNDMAN, R. J. *Forecasting methods and applications*. 3. ed. New York: John Wiley & Sons, 1998.
- MORRETIN, P. A.; TOLOI, C. M. C. *Previsão de séries temporais*. 2. ed. São Paulo: Atual Editora, 1987.
- NELSON, C. R. *Applied time series analysis for managerial forecasting*. San Francisco: Holden-Day, 1973.
- PELLEGRINI, F. R.; FOGLIATTO, F. Estudo comparativo entre modelos de Winters e de Box-Jenkins para a previsão de demanda sazonal. *Revista Produto & Produção*, v. 4, número especial, p. 72-85, abr. 2000.
- SLINI, L.; KARATZAS, K.; MOUSSIOPOULOS, N. Statistical analysis of environmental data as the basis of forecasting: an air quality application. *The Science of the Total Environment*, v. 288, p. 227-237, 2001.

## DEMAND FORECASTING: AN APPLICATION OF THE BOX-JENKINS MODELS IN THE TECHNICAL ASSISTANCE OF PERSONAL COMPUTER

### **Abstract**

*Demand forecasting is an important tool to aid on the determination of necessary resources of a given company. In this paper, the Box-Jenkins methodology was applied to analyze historical data of a personal computer repair company and provide a forecast for the number of service calls. The company studied presents three segments of clients: contracts, warranty, and on-call. As each client has its own characteristics, in order to better represent tendency and seasonality behavior through the Box-Jenkins models, a specific forecasting model was developed for each segment. The choice of the optimum models were based into graphic analysis and statistical tests, which lead to the decision of adopting the AR(1) model to foresee the number of contract clients, the ARIMA(2,1,0) model for warranty clients and the SARIMA(0,1,0)(0,1,1)<sub>12</sub> seasonal model for on-call clients.*

**Key words:** forecasting, Box-Jenkins models, ARIMA models, time series, technical assistance.