# 2
# PREVISÃO DA DEMANDA

Abandonando um pouco a visão “romântica” do termo previsão, milhares de anos após as grandes civilizações da nossa história, a “previsão do futuro” volta a tomar a sua posição de importância no plano de gestão e planejamento das empresas. Não é preciso aqui relembrar os termos mais utilizados (e os seus significados) no mundo dos negócios nos últimos anos: globalização, competitividade, *supply chain management* e redução de custos. Estes são alguns dos *drivers* que orientam as empresas na busca pelo seu maior objetivo: Maximização da Lucratividade.

Aqui pode-se introduzir o conceito de previsão de demanda, como uma das ferramentas que mais tem obtido importância nos últimos anos.

## 2.1.
## CONSIDERAÇÕES INICIAIS

O processo de previsão de demanda se baseia no pressuposto que a demanda passada de um produto poderá, em parte, explicar a sua demanda futura. Este processo busca a identificação e o entendimento dos fatores que formaram esta demanda passada e a partir daí extrapolá-los para o futuro. O conhecimento que uma empresa possui sobre o comportamento dos seus clientes é um diferencial competitivo, pois auxilia consideravelmente neste planejamento. Quanto maior a precisão na identificação destes fatores, menor será o erro associado à previsão e, portanto, melhores resultados serão alcançados. Alguns fatores podem influenciar a demanda, como por exemplo, promoções e descontos por quantidade, campanhas de vendas, lançamento de um novo produto similar, variação do clima (fator preponderante para sorvetes e cerveja, por exemplo), conjuntura econômica e relação entre produtos (promoção de um similar “roubando” demanda de outro).

Aqui abre-se um parêntese para tratar de um conceito extremamente importante apresentado por Ballou (Gerenciamento da Cadeia de Suprimentos, 2001). É o de demanda espacial. Ele destaca que, em geral, quando se fala em previsão da demanda, os gestores acabam pensando somente na questão temporal, ou seja, tratam unicamente da variação da demanda ao longo do tempo. Contudo, na Logística, além da dimensão tempo, deve-se também considerar a dimensão espacial. Em outras palavras, na logística deve existir a preocupação não somente quando a demanda irá ocorrer, mas também **aonde** ocorrerá.

A localização da demanda auxiliará ao gestor a planejar localizações de centros de distribuição, o nível de estoque no sistema logístico (CD´s x pontos de vendas), bem como dimensionamento de frota. Ainda, segundo Ballou, "Técnicas de previsão devem ser selecionadas para refletir as diferenças geográficas que podem afetar os padrões de demanda. Também, as técnicas, podem diferir dependendo se toda a demanda está prevista e, então, é desagregada por localização geográfica (previsão *top-down*) ou se cada localização geográfica está prevista separadamente e agregada posteriormente (previsão *bottom-up*)". Entretanto o presente estudo tratará somente da demanda temporal, trabalhando com demanda agregada.

Algumas considerações precisam ser levantadas sobre a questão das análises *top-down* e *bottom-up*. Cada uma das abordagens possui vantagens e desvantagens. A primeira, também denominada análise consolidada, refere-se à avaliação ou do total de uma família de produtos ou de vendas de uma diretoria regional, por exemplo. A partir da previsão consolidada, faz-se a divisão dos valores com base nos percentuais históricos de participação de cada produto ou área geográfica. Já na análise *bottom-up* parte-se de uma avaliação individualizada (produto ou região), que posteriormente é consolidada num plano de produção ou metas de vendas de uma diretoria comercial. A decisão sobre qual metodologia deve-se seguir traz implicações para a empresa. A análise consolidada é menos complexa e com menos custo, visto que não há necessidade de se analisar o comportamento de centenas (ou até mesmo, milhares de itens). Contudo, quando se avalia uma família de produtos de forma consolidada, corre-se o risco de “extrapolar” o crescimento apresentado por um determinado produto para toda a família, o que pode levar a uma distorção dos níveis de estoque.

Em resumo, a avaliação *top-down* é melhor na tentativa de se minimizar o impacto de componentes aleatórios, em contrapartida, a *bottom-up* pode ser a mais apropriada nos casos em que há presença de forte tendência (de crescimento ou declínio) em alguns itens analisados, trazendo o risco de superavaliar ou subavaliar os outros itens.

A metodologia mais adequada deve ser definida a partir da avaliação do comportamento das séries. De qualquer maneira, um dos *drivers* mais importantes e que deve ser levado em consideração no momento da decisão é o nível de precisão desejado para a previsão.



## 2.2. CARACTERÍSTICAS DA PREVISÃO

A disseminação do uso da tecnologia da informação e gerenciamento de dados tem impulsionado a utilização das técnicas de previsão de demanda. Já é possível com os sistemas integrados o acesso rápido às informações necessárias ao processo de previsão.

Esta disseminação também é conseqüência direta do desenvolvimento e popularização dos microcomputadores. Atualmente, praticamente qualquer gestor tem condições de montar modelos de previsão de demanda utilizando planilhas eletrônicas como apoio ao processo de tomada de decisão.

Contudo, apesar desta facilidade, os gestores têm que prestar atenção em algumas características das previsões:

1. Toda e qualquer previsão possui um “erro” intrínseco chamado de incerteza da demanda. Infinitos fatores podem influenciar a demanda de determinado produto, tornando impossível ao gestor mapear e avaliar o impacto de cada fator em determinado momento. O papel do gestor é

minimizar este erro. As medidas de nível de erro são fundamentais para se avaliar a qualidade da previsão;

2. As previsões de curto prazo são mais precisas que as de longo prazo, isto porque o nível de incerteza aumenta à medida que o horizonte da previsão se distancia do momento atual;
3. As previsões agregadas são mais precisas que as desagregadas. Esta característica pode ser explicada através do Teorema Central do Limite, que comprova que a soma de muitas variáveis aleatórias tende para uma variável aleatória de distribuição normal.

Outro conceito importante no processo de previsão é o de série histórica ou temporal. Segundo Chatfield ("The Analysis of Time Series", 1995) uma série histórica é uma coleção de observações realizadas seqüencialmente no tempo. A série histórica é a principal informação utilizada para se tentar entender a demanda passada. Contudo, deve-se também considerar na análise das séries temporais, que deverá existir uma dependência entre as observações. A Prof.ª Monica Barros, em seu livro "Processos Estocásticos" (2004) coloca que "a dependência serial entre os valores da série é um aspecto essencial, pois nos permite gerar previsões de valores futuros da série. Estas previsões seriam puro 'chute' se não houvesse a dependência serial". É exatamente esta dependência que reflete as informações de comportamento nos dados da série. Na figura 1, a seguir, observa-se alguns exemplos de séries históricas.

O processo de previsão de uma série temporal de demanda é composto de duas etapas:

1. Análise e modelagem da série temporal – Nesta etapa avalia-se os fatores que influenciam a demanda e suas características mais relevantes. Também busca-se representar matematicamente este comportamento;
2. Previsão – Após a etapa de análise, o modelo é aplicado e os números da previsão são gerados.

Um ponto importante sobre a avaliação das séries temporais e que é levantado em boa parte dos livros que tratam do assunto: a primeira medida a ser tomada ao se deparar com uma série histórica é construir o seu gráfico para observar a sua evolução ao longo do tempo. Chatfield (The Analysis of Time Series, 1995) é mais incisivo: “qualquer um que tente analisar uma série histórica sem ‘plotar’ os dados primeiro estará pedindo problemas. O gráfico poderá mostrar além das tendências e sazonalidades, os chamados *outliers*, ou variações anormais da demanda”. Os *outliers* necessitam ser identificados e explicados. Após a sua identificação, a variação anormal deverá ser de alguma forma ajustada, de maneira que a análise dos dados não fique distorcida.

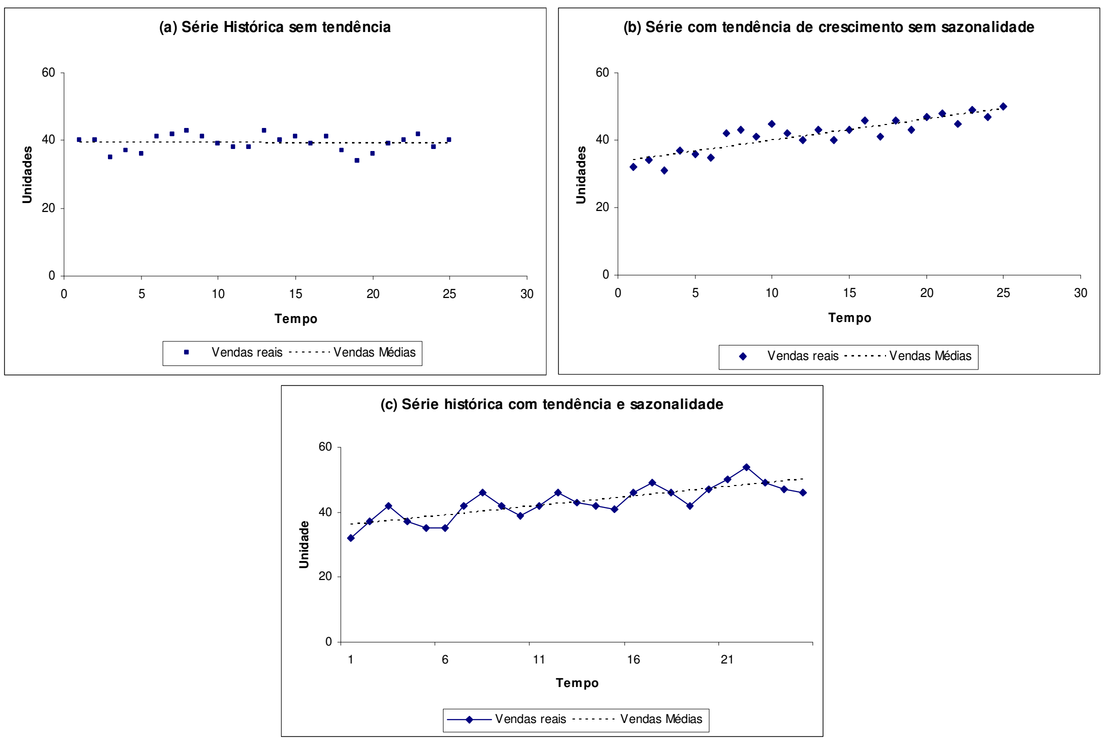


Figura 1 Séries históricas com tendência e sazonalidade.
FONTE: Adaptado de “Gerenciamento da Cadeia de Suprimentos”, Ronald Ballou, 2001.

Peter Brockwell e Richard Davis (Introduction to Time Series and Forecasting, 2002) também ressaltam a importância da montagem dos gráficos. A sua principal função é exibir a existência de tendências, sazonalidades, mudanças aparentes no comportamento da curva e possíveis *outliers*.

Os modelos de previsão podem ser classificados da seguinte maneira:

1. **Qualitativos –** Os modelos qualitativos, também denominados subjetivos, se baseiam na intuição, pesquisas de mercado ou julgamento de uma ou várias pessoas para se chegar à previsão final. Geralmente buscam-se pessoas com vasta experiência no mercado ou no assunto tratado. São muito utilizados quando não existem dados históricos disponíveis para a análise ou no lançamento de produtos novos. Alguns exemplos de métodos qualitativos que podem ser citados: estimativas da área de vendas, que levam em consideração a opinião da equipe de vendedores (é bastante útil quando esta equipe está bem próxima dos clientes); painéis de consenso, quando vários especialistas são reunidos e são incentivados a discutir as previsões; e também a analogia histórica que leva em consideração experiência anterior com produtos similares;
2. **Métodos Quantitativos ou de projeção histórica –** Referem-se aos métodos que utilizam as séries temporais para gerar a previsão. Partem do pressuposto que a demanda passada poderá auxiliar na geração de números que se aproximem da demanda futura. Conforme mencionado anteriormente, é necessária a existência de dependência entre os dados ao longo do tempo;
3. **Métodos Causais –** Procuram estabelecer uma relação entre a demanda e fatores conjunturais e/ou externos, como por exemplo, inflação, taxas de juros, políticas econômicas, temperatura ambiente.



O gestor deverá selecionar o método mais adequado às suas necessidades. Ainda segundo Chatfield, o perfil da série histórica definirá o método mais adequado. Se o perfil for bastante claro, onde tenhamos condições de verificar a existência (ou não) de tendência e sazonalidade, o método quantitativo será perfeitamente adequado às necessidades do gestor. De qualquer forma, o ideal é que a empresa busque sempre o aperfeiçoamento das previsões através da intervenção do homem. Neste momento, a composição com a análise qualitativa pode trazer excelentes resultados.

A série histórica é formada por dois componentes básicos: o **sistemático**, que se refere ao valor esperado da demanda. É composta, por sua vez, por outras três partes: <u>nível</u>, que é o comportamento das vendas atuais desazonalizada;

tendência, que significa a taxa de crescimento ou decréscimo apresentada pela série a médio e longo prazo, e sazonalidade que representa as variações que ocorrem ao longo do ano (Natal, verão, inverno, Dia das Mães, etc.).

O componente sistemático pode assumir dois comportamentos básicos:

- Aditivo: Neste caso o padrão de comportamento da sazonalidade não se modifica com possíveis alterações no nível da série;

VALOR PREVISTO (P) = NÍVEL (N) + TENDÊNCIA (T) + SAZONALIDADE (S)

- Multiplicativo: No modelo multiplicativo, a sazonalidade da série é afetada com as modificações do nível. Pode ser ampliada ou reduzida.

VALOR PREVISTO (P) = NÍVEL (N) x TENDÊNCIA (T) x SAZONALIDADE (S)

Ou

VALOR PREVISTO (P) = (NÍVEL (N) + TENDÊNCIA (T)) x SAZONALIDADE (S)

O segundo componente é o chamado **aleatório**, que se refere a todas as outras variações não explicadas pelos componentes sistemáticos. Podem ser causadas por eventos específicos e que não se repetem. Está desvinculado da parte sistemática do modelo. Pode-se afirmar que é impossível prever este último componente, cabendo à empresa avaliar o seu tamanho e variabilidade, o que gerará as chamadas medidas de erro da previsão. O papel do gestor é minimizar este erro, em outras palavras, procurar a maior precisão possível para o sistema de previsão. Em geral, supõe-se que o erro aleatório tem média zero e variância constante, e freqüentemente faz-se a hipótese de Normalidade do erro, permitindo a construção de intervalos de confiança e testes de hipóteses. Posteriormente será abordada a questão sobre indicadores de erro (também denominados de indicadores de desempenho).

VALOR PREVISTO(P)=NÍVEL(N) + TENDÊNCIA(T) + SAZONALIDADE(S) + COMPONENTE ALEATÓRIO (ε)

A figura 1, vista anteriormente, exibe três classes de modelos:

1) **Modelo Constante** – Pode ser considerado o modelo mais simples, pois não são encontrados tendências de longo prazo (crescimento ou queda) e tampouco sazonalidade. Neste caso, a demanda deverá variar somente em torno do nível. Tem-se a seguinte estrutura:

$$x_t = a + \varepsilon_t \qquad \text{eq(1)}$$

Onde define-se $a$ como o nível da demanda e $\varepsilon_t$ como o componente aleatório (ou erro), que terá média zero e variância $\sigma^2$ (figura 1.a);

2) **Modelo Linear** – Neste modelo já existe a presença da tendência (crescimento ou queda). Pode-se representá-lo com a seguinte equação geral:

$$x_t = a + b\ t + \varepsilon_t \qquad \text{eq(2)}$$

Onde denomina-se $b$ como taxa de crescimento ou de decréscimo da demanda (figura 1.b);

3) **Modelo Sazonal** – Aqui, além da tendência, passa-se a observar variações da demanda dentro do ano. Sua representação será dada por:

$$x_t = (a + b\ t)F_t + \varepsilon_t \qquad \text{eq(3)}$$

Onde $F_t$ será o fator sazonal no período $t$ (figura 1.c).

A seguir serão apresentados alguns métodos de previsão que poderão ser adotados de acordo com o perfil apresentado na série.

## 2.3. MÉTODOS DE PREVISÃO

A função dos métodos é possibilitar ao analista estimar os componentes da demanda. De acordo com os perfis apresentados no item anterior, define-se os modelos que serão mais adequados a cada série e que possibilitarão a geração de estimativas destes componentes com a menor margem de erro possível.

Atualmente existem dezenas de métodos para cálculo dos estimadores (dos mais simples aos mais complexos), contudo a referência bibliográfica consultada indica que muitas vezes os métodos 'mais simples' (portanto, com menores custos) atingem resultados tão bons quanto os mais complexos. Segundo Makridakis ("Forecasting and Planning: An Evaluation", Management Science 27, n◦ 2 – fevereiro 1981, 115-138), **em geral**, a complexidade dos modelos de previsão não diminui consideravelmente a margem de erro.

Os métodos básicos mais utilizados serão apresentados a seguir.

### 2.3.1. Média Simples

É considerado um dos métodos mais simples de previsão. Parte do pressuposto que a média aritmética simples é uma boa estimativa de previsão para os períodos seguintes:

$$P_{t+1} = \frac{\sum_{t=1}^{n} R_t}{n} \qquad \text{eq(4)}$$

Onde:

$P_{t+1}$ - Previsão para o período *t+1*;

$R_t$ - Demanda real ocorrido em *t*;

$n$ - Número total de registros ou ocorrências passadas.

Pode ser um método adequado para os modelos constantes, contudo é desaconselhável seu uso para séries com tendência e sazonalidade, visto que estes componentes provavelmente distorcerão a média, aumentando o nível de erro. Outro ponto negativo deste modelo é que a média responde muito lentamente às alterações recentes no nível da série, o que pode gerar problemas de falta de produtos, por exemplo, caso o gestor trabalhe com uma política de estoques enxuta. Uma das características deste tipo de modelo é que todas as ocorrências possuem o mesmo peso ou importância no cálculo de previsão, o que pode ser inoportuno, caso a intenção do gestor seja dar às observações mais antigas menor impacto.



### 2.3.2. Média Móvel

Trata-se de uma variação da média simples. Neste caso, passa-se a utilizar a média aritmética das $n$ últimas observações. A cada novo período substitui-se o registro mais antigo pelo mais recente e recalcula-se a média. Sua fórmula passa a ser:

$$P_{t+1} = M_t = \frac{R_t + R_{t-1} + R_{t-2} + (\ldots) + R_{t-n+1}}{n} \qquad \text{eq(5)}$$

Onde:

$P_{t+1}$ - Previsão para o período $t+1$;

$M_t$ - Média Móvel para o período $t$;

$R_t$ - Demanda real ocorrida em $t$;

$n$ - Número de registros que deverão ser considerados para o cálculo da média móvel.

Aqui, a exemplo da Média Simples, as ocorrências também possuem o mesmo peso. A diferença é que a média móvel passa a ser um pouco ‘mais sensível' às alterações do nível, já que retira as ocorrências mais antigas, substituindo-as pelas mais recentes. Estas passam a ser mais relevantes para a composição da média. Contudo, uma questão importante a ser definida é o tamanho mais adequado desta 'amostra' *n*. Segundo a Prof.ª Mônica Barros, "quanto maior o valor de *n*, mais suave será a previsão. Ao contrário, se *n* é menor, a previsão passa a oscilar muito". Logicamente, ao se adotar a amostra inteira como *n*, retorna-se à média simples. Deve-se definir *n* de modo a minimizar a soma dos quadrados dos erros da previsão. A exemplo do que ocorre com a média simples, é desaconselhável seu uso em séries temporais que apresentam tendência e sazonalidade, visto que as previsões acabam sendo reativas às alterações dos componentes da série.

### 2.3.3. Média Móvel Dupla

Este modelo já pode ser utilizado em séries com tendência. O seu cálculo é dividido em cinco etapas:

1. Cálculo da média móvel de tamanho *n*

$$M_t = \frac{R_t + R_{t-1} + R_{t-2} + (...) + R_{t-n+1}}{n} \qquad \text{eq(6)}$$

2. A partir das médias móveis apuradas no passo anterior, calculamos a média móvel dupla de tamanho *n*

$$MMD_t = \frac{M_t + M_{t-1} + M_{t-2} + (...) + M_{t-n+1}}{n} \qquad \text{eq(7)}$$

O nível é estimado somando-se a média móvel simples à diferença entre a média dupla e a simples:

$$a_t = M_t + (M_t - MMD_t) \qquad \text{eq(8)}$$

3. Para incluir a tendência no cálculo, define-se um fator de ajuste, equivalente ao coeficiente angular de reta:

$$b_t = \frac{2}{n-1}(M_t - MMD_t) \qquad \text{eq(9)}$$

4. A previsão futura será dada, então, pela equação:

$$P_{t+1} = a_t + b_t \,. \qquad \text{eq(10)}$$

Da mesma maneira que na média móvel, também a necessidade de se definir a quantidade de períodos *n* mais adequada à série em questão (sempre buscando minimizar a soma quadrada dos erros da previsão do período seguinte). Aqui as ocorrências também irão possuir o mesmo peso, independentemente da sua "idade", o que é um dos pontos negativos deste tipo de modelo.

A seguir, na tabela 1 e na figura 2, observa-se um comparativo dos métodos de médias, simples, móveis e duplas com tamanho três (n=3), utilizando os dados de estojos de meia aliança.

| DESCRICAO | MES | 2004 | média simples | média móvel n=3 | média dupla n=3 |
|---|---|---|---|---|---|
| ESTOJO ANEL/MEIA ALIANCA | 1 | 1448 | 1720 | | |
| | 2 | 1836 | 1720 | | |
| | 3 | 1822 | 1720 | 1702 | |
| | 4 | 1638 | 1720 | 1765 | |
| | 5 | 1715 | 1720 | 1725 | 1731 |
| | 6 | 1380 | 1720 | 1578 | 1689 |
| | 7 | 1220 | 1720 | 1438 | 1580 |
| | 8 | 1591 | 1720 | 1397 | 1471 |
| | 9 | 1471 | 1720 | 1427 | 1421 |
| | 10 | 1819 | 1720 | 1627 | 1484 |
| | 11 | 2120 | 1720 | 1803 | 1619 |
| | 12 | 2579 | 1720 | 2173 | 1868 |

Tabela 1 Comparativo dos resultados das técnicas de média simples, móvel e dupla, em unidades.

FONTE: Gerência da CAM

Neste caso, não se buscará a otimização do tamanho da janela (*n)*, que será fixado $n = 3$.

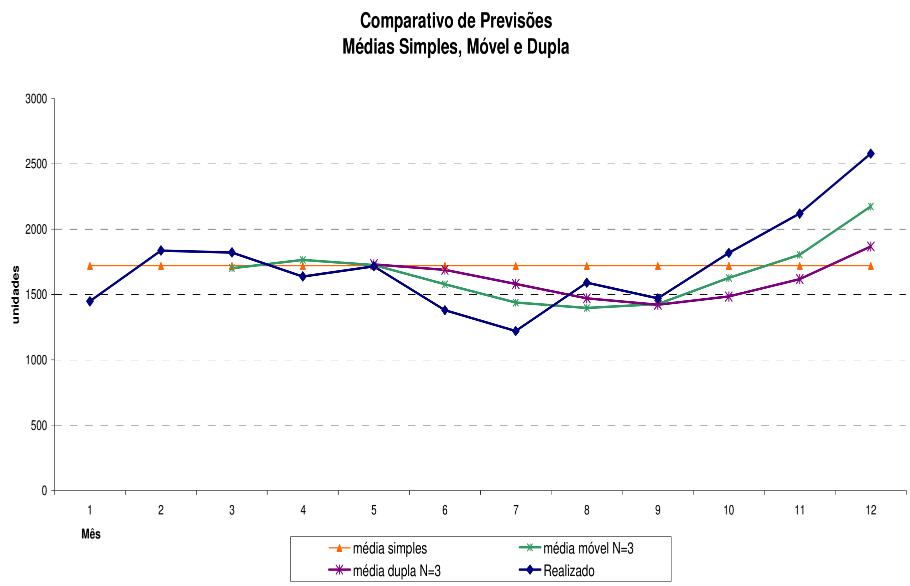


Figura 2 Comparativo de resultados de previsões, utilizando a média simples, móvel e dupla.

FONTE: Gerência da CAM.

Neste caso, se houvesse sido adotado o modelo de média simples, poderia-se considerar a cobertura adequada nos primeiros cinco meses do ano, contudo a partir do sexto até o sétimo mês, há uma queda no consumo dos estojos, tendência que não é acompanhada pela média simples, que é fixa. A média móvel responde um pouco melhor a esta alteração, começando a "reagir" a partir do mês seguinte. A partir do oitavo mês, a demanda começa a apresentar um crescimento acentuado. Este movimento começa a ser acompanhado pelos números gerados pela média móvel somente a partir do décimo mês, mas sem a mesma "velocidade" da taxa de crescimento. Ou seja, em termos práticos, a partir do décimo mês o risco de desabastecimento aumenta.

A média móvel dupla apresenta, neste caso, uma previsão mais próxima da demanda ao longo do tempo.  Até o décimo mês podemos observar que este método apresenta erros menores que os outros, contudo a sua resposta ao aumento de demanda a partir do décimo mês, acaba ficando abaixo da média móvel. A partir deste momento a Média Móvel passa a estar mais perto do realizado que a Média Móvel Dupla.

### 2.3.4. Amortecimento Exponencial Simples

O método de Amortecimento Exponencial é um método similar à média móvel, mas que tem como base a diferenciação dos pesos aplicados aos valores da série histórica em função do tempo. As observações mais recentes da série terão uma importância maior na obtenção da previsão que os primeiros eventos. O conceito por trás deste modelo pressupõe que os últimos eventos de uma série terão "informações mais atualizadas" sobre a demanda e, conseqüentemente, mais importância na previsão da demanda futura. Os diversos métodos de amortecimento exponencial são considerados como alguns dos mais úteis, pela sua simplicidade e pequena necessidade de séries longas, o que facilita a implementação. Na literatura podem ser encontradas diversas variações dos métodos de amortecimento, mas sempre mantendo o mesmo conceito chave: ponderação das observações em função do tempo. Algumas metodologias de amortecimento serão apresentadas neste trabalho.

Os pesos são conhecidos como fatores de ponderação ou coeficientes de amortecimento e seus valores variam de zero a um. Segundo Barros (2004), "A existência de uma ou mais constantes de amortecimento irá determinar como funciona este mecanismo de ponderação, ou o quão rapidamente decai a influência das observações passadas". Esta ponderação faz com que o modelo responda mais rapidamente às mudanças na série.

O método é denominado exponencial, pois os pesos ($\alpha$) decrescem exponencialmente, como uma progressão geométrica ($\alpha$, $\alpha(1-\alpha)$, $\alpha(1-\alpha)^2$, $\alpha(1-\alpha)^3$), à medida que os eventos vão ocorrendo ao longo do tempo.

Quanto maior o valor de $\alpha$, maior a influência do evento mais recente na previsão final. Contudo, um valor de $\alpha$ muito elevado pode fazer com que a previsão fique extremamente sensível às variações aleatórias da demanda realizada, em detrimento das mudanças estruturais. Por outro lado, um coeficiente muito baixo tornará o modelo extremamente estável, resistente à aleatoriedade existente na série histórica. A bibliografia relacionada ao assunto sugere um valor de $\alpha$ entre 0,05 e 0,25.

Neste método, a maior dificuldade é exatamente definir o valor adequado de $\alpha$, de forma a minimizar os erros da previsão, buscando ao mesmo tempo uma aderência mais rápida às mudanças ocorridas na série histórica. Uma possibilidade é a escolha de $\alpha$ que minimize o erro quadrático médio (EMQ), que é uma estimativa da variância do erro da previsão.

$$EMQ = \frac{1}{n}\sum_{t=1}^{n}(R_t - P_t)^2 \qquad \text{eq(11)}$$

Onde:

$R_t$ - Valor observado no instante *t*;

$P_t$ - Previsão no mesmo instante.

Atualmente já existem softwares de séries temporais que estimam os coeficientes ótimos de amortecimento. Até mesmo o *Excel* pode auxiliar nesta tarefa, através da ferramenta *Solver*.

Outro ponto importante no amortecimento exponencial é o valor inicial a ser adotado na previsão, pois dependendo do tamanho da série, aquele terá um peso muito grande na estimação da previsão. Geralmente adota-se $P_0 = R_0$. O efeito da escolha inicial é minimizado, é claro, à medida que *t* aumenta. Uma alternativa bastante utilizada nesta questão é trabalhar com a média de *n* valores iniciais como ponto de partida.

O amortecimento exponencial simples é um método que, a exemplo da média móvel simples, não é indicada para séries temporais com tendência e sazonalidade, pois nesta situação as previsões geradas acompanhariam a evolução da demanda com atraso.

A equação de previsão no modelo de amortecimento exponencial simples é dado por:

$$P_{t+1} = \alpha R_t + \alpha(1-\alpha)\ R_{t-1} + \alpha(1-\alpha)^2\ R_{t-2} + (...) \qquad \text{eq(12)}$$

Onde:

$P_{t+1}$ - Previsão para o período *t+1*;

$\alpha$ - Coeficiente de amortecimento, $\alpha \in [0,1]$;

$R_t$ - Demanda real ocorrido em *t*;

Ou também, na forma reduzida: $P_{t+1} = \alpha R_t + (1-\alpha)P_t$ eq(13)

A exemplo do que ocorre com qualquer outro método de previsão, eventualmente observa-se a necessidade de revisão de $\alpha$, que será definida através do monitoramento dos níveis de erro das previsões. É importante a definição de políticas de controles para estes erros com os respectivos ajustes.

### 2.3.5. Amortecimento Exponencial Duplo (Modelo De Brown)

O modelo de Brown é uma variação do amortecimento exponencial simples desenvolvida para séries temporais que apresentam tendência, mas não sazonalidade. Este método se assemelha à Média Móvel Dupla, sendo o componente de nível calculado através da diferença entre um valor apurado com o amortecimento simples e outro, com um "segundo amortecimento" do primeiro. Já para o cálculo da tendência, existe uma equação própria, similar à fórmula de coeficiente angular da reta. O cálculo de previsão ocorre como a seguir:

1. Cálculo do primeiro amortecimento: $P_t = \alpha R_t + (1-\alpha)P_{t-1}$ eq(14)

2. Cálculo do segundo amortecimento: $P_t^{'} = \alpha P_t + (1-\alpha)P_{t-1}^{'}$ eq(15)

3. Cálculo de estimativa do nível como a soma do primeiro amortecimento com a diferença entre o primeiro e o segundo amortecimentos:

$$a_t = P_t + (P_t - P_t^{'}) = 2P_t - P_t^{'} \qquad \text{eq(16)}$$

4. Cálculo de estimativa de tendência: $b_t = \frac{\alpha}{1-\alpha}(P_t - P_t^{'})$ eq(17)

5. A previsão futura será dada, então, pela equação: $P_{t+p} = a_t + b_t p$ eq(18)

Aqui novamente deve-se destacar que tão importante quanto a constante de amortecimento a ser utilizada, são as estimativas iniciais de $P_0$ e $P_0^{'}$. Não se deve utilizar o primeiro valor da série para iniciar as equações, pois corre-se o risco de subestimar a tendência desta série. As estimativas iniciais do $P_0$ e $P_0^{'}$ são geralmente dadas por:

$$P_0 = a_0 - \frac{1-\alpha}{\alpha} b_0 \quad \text{eq(19)}$$

$$P_0^{'} = a_0 - 2\frac{1-\alpha}{\alpha} b_0 \quad \text{eq(20)}$$

Onde:

$a_0$ = coeficiente linear obtido através da regressão linear dos valores da série pelos números dos períodos;

$b_0$ = coeficiente angular obtido através da regressão dos valores da série pelos números dos períodos.

Para a estimativa do valor inicial de $\alpha$, necessário para o cálculo das estimativas iniciais do $P_0$ e $P_0^{'}$, deve-se atuar da mesma forma que a descrita no método de amortecimento exponencial simples, quando mencionou-se que $\alpha$ é o valor que deve ser escolhido de forma a minimizar o erro médio quadrado (EMQ).

### 2.3.6. Amortecimento Exponencial - Método De Holt

O método de Holt é outra variação do método de amortecimento, utilizado para as séries que apresentam tendência (mas não sazonalidade). Neste método, ao contrário do modelo de Brown, adotam-se duas constantes de amortecimento diferentes, $\alpha$ e $\beta$, uma delas para estimação da tendência e a outra para o nível. Aqui recuperaremos o conceito de componente sistemático, apresentado no início deste capítulo.

| VALOR PREVISTO (P) = NÍVEL (N) + TENDÊNCIA (T) |
|---|

O método de Holt trabalha com três equações:

1. Estimação do nível: $N_t = \alpha R_t + (1-\alpha)(N_{t-1} + T_{t-1})$ eq(21)
2. Estimação da tendência: $T_t = \beta(N_t - N_{t-1}) + (1-\beta)T_{t-1}$ eq(22)
3. Equação para previsão para *p* períodos futuros: $P_{t+p} = N_t + T_t p$ eq(23)

Onde:

$N_t$ - Nível para o período *t;*

$R_t$ - Demanda real ocorrido em *t*;

$T_t$ - Tendência para o período *t*;

$P_{t+1}$ - Previsão para o período *t+1*;

$\alpha$ - Coeficiente de amortecimento para o nível, $\alpha \in [0,1]$;

$\beta$ - Coeficiente de amortecimento para a tendência, $\beta \in [0,1]$;

A exemplo do que ocorre nos outros métodos de amortecimento exponencial, é necessário calcular as estimativas iniciais para o nível e a tendência de forma a iniciar o sistema de cálculo das previsões. Neste caso, pode-se trabalhar com a regressão linear entre a demanda e os períodos, pois existe uma tendência (de crescimento ou declínio) ao longo do tempo. A estimativa para o nível inicial $N_0$ será, então, o coeficiente linear obtido através da regressão linear da demanda pelo número de período. E a tendência $T_0$, será o coeficiente angular ou de inclinação, obtido através do mesmo processo de regressão.

Cabe aqui destacar que o Excel possui uma ferramenta que realiza facilmente o cálculo de regressão linear (menu **Ferramentas – Análise de Dados – Regressão**).

### 2.3.7. Amortecimento Exponencial – Modelo de Winters

O modelo de Winters é utilizado para as séries que apresentam tendência e sazonalidade e adota três coeficientes de amortecimento diferentes, um para cada componente sistemático da demanda (Nível, Tendência e Sazonalidade), $\alpha$, $\beta$ e $\gamma$, respectivamente, todos no intervalo (0,1)

Novamente, resgatamos o conceito de componente sistemático, agregando a sazonalidade. Pode-se ter dois comportamentos básicos:

- Aditivo: O padrão de comportamento da sazonalidade não se modifica com possíveis alterações no nível da série

| |
|---|
| VALOR PREVISTO (P) = NÍVEL (N) + TENDÊNCIA (T) + SAZONALIDADE (S) |

- Multiplicativo: No modelo multiplicativo, a sazonalidade da série é afetada com as modificações do nível. Pode ser ampliada ou reduzida.

| |
|---|
| VALOR PREVISTO (P) = NÍVEL (N) x TENDÊNCIA (T) x SAZONALIDADE (S) |

Ou

| |
|---|
| VALOR PREVISTO (P)= (NÍVEL (N) + TENDÊNCIA (T))x SAZONALIDADE (S) |

O gestor, então, terá que avaliar e selecionar o comportamento com base na sua série histórica. No presente estudo, será utilizado o modelo multiplicativo (também chamado de misto) para demonstrar o método.

O método de Winters multiplicativo trabalha com quatro equações:

1. Estimação do nível: $N_t = \alpha(\frac{R_t}{S_t} + (1-\alpha)(N_{t-1} + T_{t-1})$ eq(24)

2. Estimação da tendência: $T_t = \beta(N_t - N_{t-1}) + (1-\beta)T_{t-1}$ eq(25)

3. Estimação da sazonalidade: $S_t = \gamma(\frac{R_t}{N_t}) + (1-\gamma)S_{t-p}$ eq(26)

   Sendo *p* o número de períodos sazonais (doze no caso de séries mensais e quatro para séries trimestrais);

4. Equação de previsão para *k* períodos futuros:

   $P_{t+k} = (N_t + T_t k)S_t$ eq(27)

Onde:

$N_t$ - Nível para o período *t;*

$R_t$ - Demanda real ocorrido em *t*;

$T_t$ - Tendência para o período *t*;

$S_t$ - Sazonalidade para o período *t*;

$P_{t+1}$ - Previsão para o período *t+1*;

$\alpha$ - Coeficiente de amortecimento para o nível, sendo que $\alpha \in [0,1]$;

$\beta$ - Coeficiente de amortecimento para a tendência, sendo que $\beta \in [0,1]$;

$\gamma$ - Coeficiente de amortecimento para a sazonalidade, sendo que $\beta \in [0,1]$;

O método é iniciado com o cálculo das estimativas iniciais para o nível, tendência e fatores sazonais. Antes de se estimar N e T, deve-se dessazonalizar a série, ou seja, remover os efeitos da sazonalidade dos números. Aqui se introduz o conceito de **período sazonal**, que passará a ser representar por ***p***. Por exemplo, numa série trimestral, existem quatro períodos sazonais por ano, ou seja, p = 4. Em séries mensais, como as consideradas neste trabalho, p = 12.

Para a dessazonalização da série deve-se calcular uma média dos períodos consecutivos da demanda *p*. No caso da periodicidade *p* ser ímpar, o cálculo fornecerá a demanda dessazonalizada para um período existente. A demanda dessazonalizada será dada, então, pela seguinte fórmula:

$$\overline{R}_t = \sum_{i=t-(p/2)}^{t+(p/2)} R_i / p \qquad \text{eq(28)}$$

Contudo se *p* for par, esta média fornecerá uma demanda dessazonalizada para um ponto entre dois períodos. Para resolver esta questão utiliza-se a seguinte fórmula:

$$\overline{R_t} = \left[ R_{t-(p/2)} + R_{t+(p/2)} + \sum_{i=t+1-(p/2)}^{t-1+(p/2)} 2R_i \right] / 2p \qquad \text{eq(29)}$$

Aplicando as fórmulas, obtem-se as demandas dessazonalizadas para os períodos da série histórica. Com esta nova demanda calculada, tem-se condições de avaliar de forma adequada a tendência da série, pois existe uma relação linear entre esta demanda dessazonalizada e os períodos. Esta relação pode ser representada pela fórmula: $\overline{R}_t = N_t + tT_t$ eq(30)

Assim, através da regressão linear são definidos os valores iniciais do nível ($N_0$) e da tendência ($T_0$). O nível será obtido pelo intercepto e a tendência pelo coeficiente angular (ou inclinação).

Ao contrário do método de Holt, não se pode aplicar a regressão linear diretamente na demanda original, pois a série tem "embutida" em si a sazonalidade, o que distorceria o resultado da regressão.

A seguir calcula-se, então, os fatores sazonais (S) para cada período da série histórica. Para tal, basta dividir a demanda original pela demanda dessazonalizada: $\overline{S}_t = R_t / \overline{R}_t$. eq(31)

A partir da periodicidade ***p*** são estimados os valores iniciais dos fatores de sazonalidade para cada período sazonal ($S_t$), através da média simples dos períodos que o compõem.

Por exemplo, dada uma série histórica com um período sazonal de três meses de um total de doze, o que significa quatro períodos de sazonalidade anuais, obtem-se os fatores de sazonalidade da seguinte forma:

$$S_1 = (\overline{S}_1 + \overline{S}_5 + \overline{S}_9)/3 \qquad \text{eq(32)}$$

$$S_2 = (\overline{S}_2 + \overline{S}_6 + \overline{S}_{10})/3 \qquad \text{eq(33)}$$

$$S_3 = (\overline{S}_3 + \overline{S}_7 + \overline{S}_{11})/3 \qquad \text{eq(34)}$$

$$S_4 = (\overline{S}_4 + \overline{S}_8 + \overline{S}_{12})/3 \qquad \text{eq(35)}$$

Cabe aqui destacar que no caso de sazonalidade multiplicativa, a soma dos fatores sazonais é igual a S (por exemplo, no caso de dados mensais, S = 12). Já na sazonalidade aditiva, a soma dos fatores sazonais é sempre igual a zero.

A partir das constantes de amortecimento ($\alpha$, $\beta$ e $\gamma$), dos valores iniciais do nível (N), tendência (T) e sazonalidade (S) e das fórmulas apresentadas pode-se atualizar as estimativas ao longo do tempo e aplicar a equação de previsão. Os fatores sazonais serão atualizados somente cada vez que o período sazonal for completo.

### 2.3.8. Modelos ARIMA de BOX e JENKINS

Os modelos de Box-Jenkins também conhecidos como ARIMA, sigla para *Auto Regressive Integrated Moving Averages* (auto-regressivos integrados de médias móveis), buscam explicar o comportamento das séries através da autocorrelação temporal entre os seus valores. O ponto chave do processo é a busca da estrutura de correlação que melhor represente este comportamento.

Os modelos ARIMA são formados por três filtros: o componente auto-regressivo (AR), o filtro de integração (I) e as médias móveis (MA).

O componente AR indica quão relacionados estão os valores sucessivos da série analisada (defasagens temporais ou *"time lags")*. A identificação do número de termos necessários é feita através da análise dos coeficientes de autocorrelação parcial. O filtro de médias móveis (MA) busca a mesma relação, só que entre os erros da série.

Aqui serão introduzidos alguns conceitos importantes para a aplicação dos modelos ARIMA:

1. Estacionariedade: Uma série é dita estacionária quando está em "equilíbrio", o que significa que tanto a média quanto a variância são constantes ao longo da série e a sua covariância é função da defasagem entre os instantes (ou *lags)*. Desta forma, como primeiro passo para se iniciar a modelagem Box-Jenkins, deve-se transformar a série para torná-la estacionária. O filtro de integração refere-se exatamente ao processo de modificação da série original através de diferenças sucessivas até que ela esteja em equilíbrio ou estacionária, de forma a permitir a identificação de todos os seus padrões de comportamento;

2. Autocovariância de *lag* k: Segundo Souza (1991) é a covariância entre $X_t$ e o seu valor $X_{t+k}$, separado por k intervalos de tempo. Sendo { $X_t$ :t=1,2,3,4,...,T}, podemos defini-la como:

$$\gamma_k = \frac{1}{T-k-1}\sum_{t=k+1}^{T}(X_t - \mu)(X_{t+K} - \mu) \qquad \text{eq(36)}$$

3. Autocorrelação de *lag* k (ACF): Segundo o mesmo autor, "a autocorrelação serve para medir a extensão para a qual o valor tomado no tempo t é dependente do realizado no tempo t-k". O gráfico da autocorrelação ao longo de k é denominado de correlograma e é fundamental para identificação da estrutura de um processo ARIMA.

Souza (1991) ainda comenta que para se obter uma boa estimativa da função de autocorrelação, necessitamos de pelo menos 50 observações. A sua fórmula é dada por:

$$\rho_k = \frac{\sum_{t=k+1}^{T}(X_t - \mu)(X_{t-k} - \mu)}{\sum_{t=k+1}^{T}(X_t - X_{t-K})^2} \qquad \text{eq(37)}$$

4. Autocorrelação Parcial (PACF): Segundo Souza (1991), é a medida de correlação entre duas observações, eliminando-se a dependência existente entre as observações intermediárias. Considerando-se quatro observações $X_1$, $X_2$, $X_3$ e $X_4$, denomina-se correlação parcial quando se deseja saber a correlação entre $X_2$ e $X_4$, eliminando-se a influência de $X_3$, por exemplo. O gráfico de PACF também é importante para a análise da estrutura destes modelos;

5. Ruído branco: É uma seqüência de valores aleatórios independentes e identicamente distribuídos. Em geral, supõe-se que eles apresentem distribuição normal com média zero, variância constante e função de autocorrelação nula em todos os *lags* de tempo;

6. Operador de atraso: Representa uma defasagem de k períodos de tempo para trás. Sua notação é dada por:

$$B^k X_t = X_{t-k} \qquad \text{eq(38)}$$

Outro conceito não menos importante é o de modelo linear. Segundo Souza (1991), busca-se através da análise de um conjunto de observações, com duas ou mais variáveis, explicar o comportamento de uma delas em função de outra. Ora, se as observações estão indexadas no tempo, através da aplicação da Teoria de Funções de Transferência em sistemas lineares, conseguimos converter uma série "input" - ruído branco - em outra "output" - processo estocástico estacionário. A figura 3, a seguir, representa o processo.


Figura 3 Representação esquemática de um modelo linear.
FONTE: Metodologia Box & Jenkins para Séries Temporais (1991)

Assim, segundo Souza (1991), "uma classe muito importante de processos estocásticos, $X_t$, é conseqüência da passagem de um processo ruído branco através de um filtro linear. Freqüentemente a filtragem linear é apenas uma combinação linear de observações passadas e presentes da série original." Matematicamente teremos:

$$X_t = \sum_{k=0}^{\infty} \psi_k a_{t-k} \qquad \text{eq(39)}$$

Onde:

- $\psi_k$ é o filtro linear definido por: $\psi_k = \dfrac{\theta_q(B)}{\phi_p(B)}$ ;

- $a_{t-k}$ - Ruído branco ou resíduo.

Desta forma, os modelos ARMA(p,q) de Box & Jenkins podem ser escritos como:

$$\underbrace{\phi_p(B)X_t}_{AR} = \underbrace{\theta_q(B)a_t}_{MA} \qquad \text{eq(40)}$$

Onde:

B - operador de defasagem com notação $B^k X_t = X_{t-k}$ ;

$\phi_p$ - Polinômio de grau p, denominado AR(p);

$\theta_q$ - Polinômio de grau q, denominado MA(q).

Os modelos são conseqüência direta da interação dos três filtros conceituados anteriormente ou de parte deles. A etapa de identificação da estrutura dos modelos é complexa, mas atualmente já existem softwares que auxiliam na modelagem do processo.

Ate o momento partiu-se do pressuposto que as séries temporais eram estacionárias. Contudo, eventualmente pode-se deparar com séries não-estacionárias, sendo necessário, então, transformá-las. A transformação mais comum se dá através do cálculo de diferenças sucessivas entre os termos da série, até que se obtenha uma série estacionária.

A primeira diferença de $X_t$ será definida por:

$$\nabla^d X_t = X_t - X_{t-1} \qquad \text{eq(41)}$$

Este termo é também denominado operador de diferença. Caso, a série permaneça não estacionária, deve-se proceder com a segunda diferença:

$$\nabla^d \ [\mathrm{D}\ X_t] = \mathrm{D}\ [X_t - X_{t-1}] = X_t - 2X_{t-1} - X_{t-2} \qquad \text{eq(42)}$$

Segundo Barros (2004): “genericamente, a série final estacionária será obtida pela aplicação de “d” (d=0,1,2,3,...) diferenças na série original. Sua notação será dada por $W_t = \nabla^d X_t$. Em seguida, modela-se a série estacionária obtida, utilizando-se a modelagem ARMA (p,q) apropriada.”

Os modelos são indicados por ARIMA (p,d,q), onde os parâmetros do modelo são: ordem do polinômio auto-regressivo (p), número de diferenças (d) e ordem do polinômio de médias móveis (q).

### 2.3.8.1.
### Modelo Auto-regressivo (AR)

Conforme mencionado anteriormente, no modelo AR a série das observações históricas $X_t$ é descrita pelos seus valores passados regredidos e pelo ruído aleatório $\varepsilon_t$. O modelo auto-regressivo de ordem p pode ser descrito como:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + (\ldots) + \phi_p X_{t-p} + \varepsilon_t \qquad \text{eq(43)}$$

Onde:

$X_t$ - Valor da série temporal no instante t;

$\varepsilon_t$ - Resíduo ou erro aleatório;

$\phi_p$ - Parâmetro que descreve como $X_t$ se relaciona com $X_{t-p}$

O modelo AR(p) pode ser reescrito utilizando-se o operador de defasagem B, assim:

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p) X_t = \varepsilon_t \Rightarrow \phi(B) X_t = \varepsilon_t \qquad \text{eq(44)}$$

A especificação do grau do polinômio *p*, que irá determinar o número de termos da equação geral, passa pela análise dos coeficientes de autocorrelação e autocorrelação parcial. Segundo Barros (2004), “num processo AR(p), as autocorrelações parciais são nulas para todos os *lags* maiores que p.” Assim, numa representação gráfica do PACF pode-se identificar o grau do modelo AR(p). A seguir, na figura 4, veremos um exemplo de gráfico do PACF, indicando um modelo AR(1).

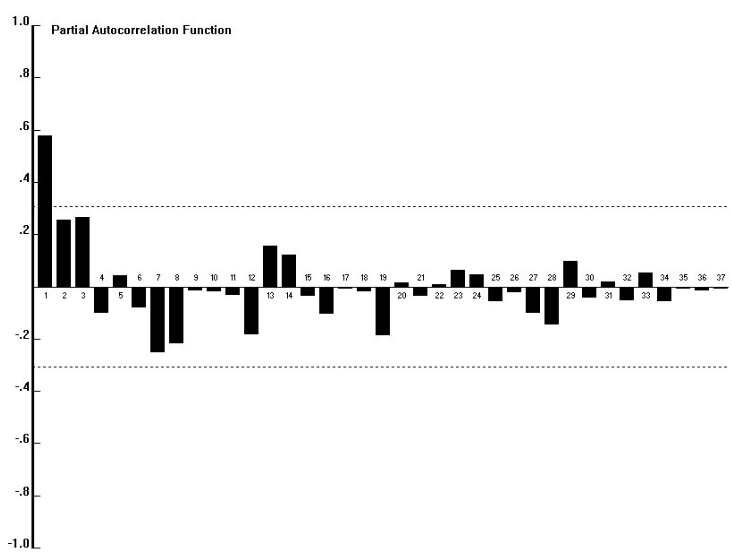


Figura 4 Gráfico PACF de uma série, indicando um modelo AR(1).

Um modelo AR(1) terá a seguinte forma: $X_t = \phi_1 X_{t-1} + \varepsilon_t$ eq(45)

Para o modelo ser considerado estacionário, se faz necessário que $|\phi_1| < 1$ e que as autocovariâncias sejam independentes.

Já um modelo AR(2), assumirá a forma:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t \quad \text{eq(46)}$$

Após definido o grau da equação, tomando como exemplo uma equação AR(2), deveremos estimar os valores de $\phi_1$ e $\phi_2$. O critério é buscar a minimização do MSE.

Assim, rearranjando a equação de AR(2), obtém-se:

$$\varepsilon_t = X_t - (\phi_1 X_{t-1} + \phi_2 X_{t-2}) \quad \text{eq(47)}$$

Sendo que deve-se minimizar $\frac{\sum \varepsilon_i^2}{n}$.

### 2.3.8.2. Modelo de Médias Móveis (MA)

Segundo Makridakis (1978):

*"Existem alguns padrões de comportamento das séries temporais que não serão identificados e isolados pelos modelos AR(p). Contudo, existe outro tipo de modelo, chamado de médias móveis, ou moving average, MA(q), que poderá auxiliar na identificação destes padrões. Qualquer série temporal discreta pode ser expressa como um modelo AR e/ou MA. Enquanto os modelos AR(q) buscam as previsões da série com base na combinação linear dos seus (p) valores passados, os modelos MA(q) realizam as previsões com base na combinação linear dos erros passados."*

O modelo de médias móveis MA(q) pode ser descrito, na sua forma geral, como:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - (...) - \theta_q \varepsilon_{t-q} \qquad \text{eq(48)}$$

Onde:

$X_t$ - Valor da série temporal no instante t;

$\varepsilon_t$ - Resíduo ou erro aleatório no instante t.

$\theta_p$ - Parâmetro de média móvel que descreve como $X_t$ se relaciona com $X_{t-p}$

O modelo de médias móveis de ordem q, ou MA(q), pode ser reescrito utilizando-se o operador de defasagem B, assim:

$$(1 - \theta_1 B - \theta_2 B^2 - ... - \theta_p B^p)\varepsilon_t = X_t \Rightarrow \theta(B)\varepsilon_t = X_t \qquad \text{eq(49)}$$

A exemplo do que ocorre com a classe de modelos AR(p), a identificação adequada do valor de q passa pela análise dos coeficientes de autocorrelação e autocorrelação parcial da série. O modelo mais simples de médias móveis é o de grau um, ou MA (1), representado pela fórmula:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} \qquad \text{eq(50)}$$

O modelo MA(2) é representado por:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \qquad \text{eq(51)}$$

O grau $q$ deve ser estimado buscando-se também a minimização do MSE.

### 2.3.8.3. Modelos Auto-regressivos de Médias Móveis (ARMA)

Em alguns casos é necessário combinar as estruturas AR(p) e MA(q), gerando um modelo ARMA (p,q). Segundo Makridakis, o desempenho combinado dos dois modelos será melhor do que se for trabalhado com o modelo AR(p) e MA(q) separadamente, resultando em modelos mais parcimoniosos. A expressão geral do modelo é dada por:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + (\ldots) + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - (\ldots) - \theta_q \varepsilon_{t-q} \quad \text{eq(52)}$$

O modelo ARMA(1, 1) é descrito como:

$$X_t = \phi_1 X_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad \text{eq(53)}$$

O modelo ARMA(2, 1) é descrito como:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad \text{eq(54)}$$

Segundo Makridakis, para uma série estacionária a metodologia proposta por Box e Jenkins pode ser resumida em três fases distintas:

1. **Identificação:** O objetivo desta fase é selecionar o modelo ARMA(p,q) mais adequado para se descrever os padrões de comportamento da série histórica. Aqui serão analisados os dados históricos e os comportamentos das funções de autocorrelação (ACF) e autocorrelação parcial (PACF), identificando desta forma os graus (p,q) do modelo;
2. **Estimação e teste:** A partir do momento em que os valores (p, q), são identificados, estima-se os parâmetros $\phi$ do modelo AR, $\theta$ do componente MA e a variância dos resíduos. A partir da análise da ACF e PACF destes resíduos tem-se condições de avaliar se o modelo é adequado ou não. Para isto, estas correlações deverão ser insignificantes, o que indicaria que o modelo está representando adequadamente o padrão de comportamento da série. No caso do modelo não ser adequado, o ciclo é repetido, retornando-se à fase de identificação. Alguns autores sugerem a identificação de mais de um modelo simultaneamente, que também serão estimados e verificados.

3. **Previsão:** Nesta etapa tem-se a aplicação do modelo para a previsão dos números futuros.

Cabe ressaltar que a metodologia sugerida acima pressupõe que a série estudada é estacionária. Se a série original não for estacionária, diferenças sucessivas deverão ser previamente aplicadas de forma a torná-la estacionária. Em outras palavras, a ordem da definição (*d*), já foi identificada.

A figura 5 ilustra de forma resumida o processo.

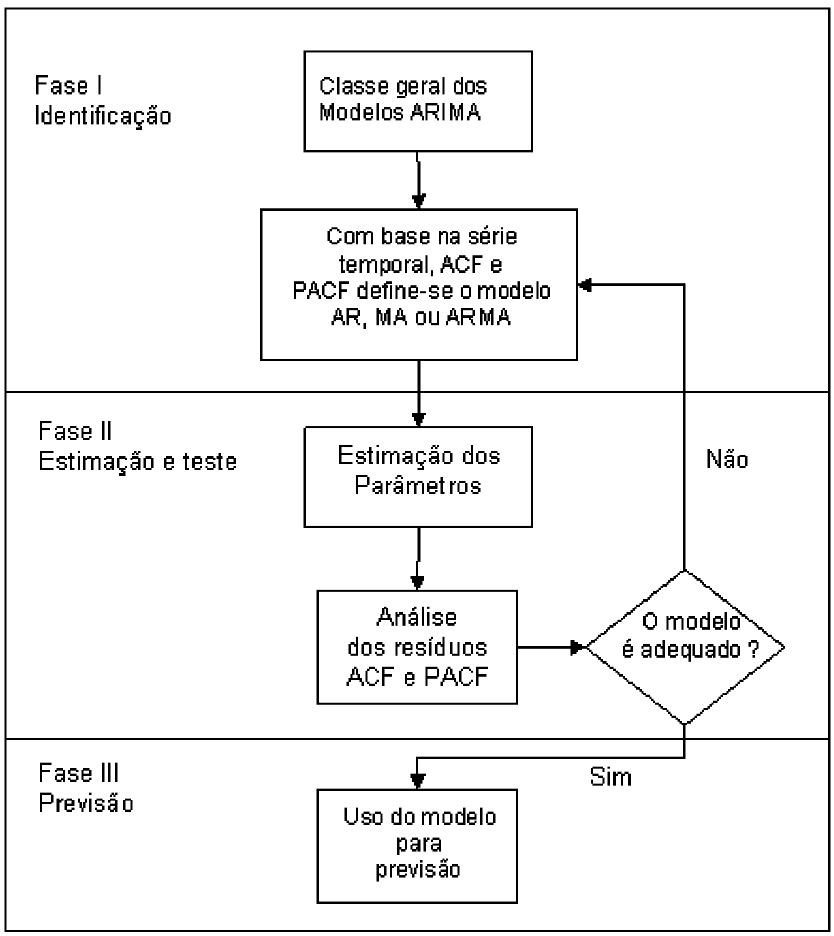


Figura 5 Abordagem do método de Box-Jenkins

FONTE: Adaptado do livro Forecasting Methods and Applications, Makridakis 1978.

Segundo Granger e Newbold (Forecasting Economic Time Series, 1986), os modelos de Box-Jenkins necessitam de uma série temporal de cinco a dez anos de dados históricos mensais (entre sessenta a cento e vinte observações) para produzir análises com precisão adequada. Makridakis (Forecasting Methods and Applications, 1978) sugere que, no mínimo, de quarenta a cinqüenta observações são necessárias para se avaliar adequadamente os padrões existentes numa série temporal.

### 2.3.8.4. Identificação

Na descrição dos modelos, mencionou-se que a identificação do grau, ou valores dos termos q e p, são feitos através da análise dos coeficientes de autocorrelação (ACF) e auto-correlação parcial (PACF) da série temporal. A identificação dos modelos pode ser feita através da observação destes gráficos. Neste tópico, serão demonstrados os comportamentos teóricos de ACF e PACF para os modelos AR e MA.

Será apresentada aqui uma abordagem superficial, baseada em Makridakis, mas suficiente para uma primeira abordagem do assunto.

O modelo AR(p) é identificado examinando-se o gráfico da função PACF. Neste caso o gráfico deverá cair abruptamente após algumas observações. Assim, o grau deste modelo será indicado pelos primeiros *lags* significativamente diferentes de zero. Já a sua função ACF apresenta uma queda tendendo a zero através de uma curva exponencial (vide figuras 6 e 7, a seguir).

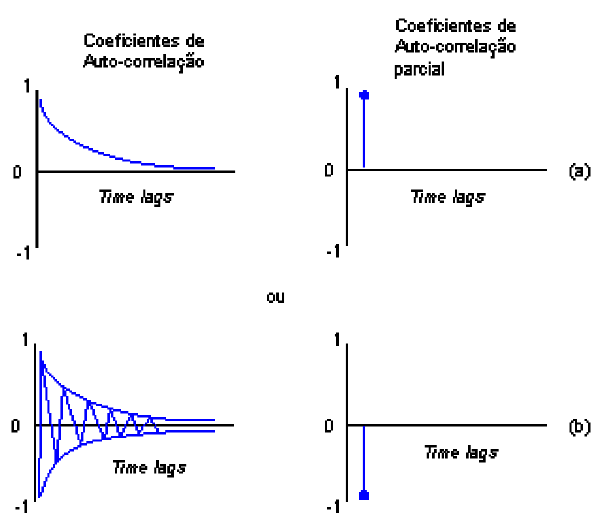

Figura 6 Gráficos Teóricos de ACF e PACF de uma série, indicando um modelo AR(1).
FONTE: Adaptado do livro Forecasting Methods and Applications, Makridakis 1978.

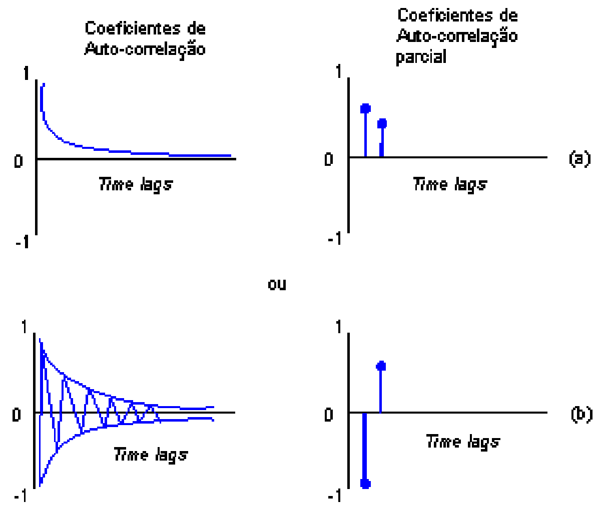


Figura 7 Gráficos Teóricos de ACF e PACF de uma série, indicando um modelo AR(2).
FONTE: Adaptado do livro Forecasting Methods and Applications, Makridakis 1978.

Com o modelo MA(q) ocorre exatamente ao contrário, no gráfico de ACF, os valores dos coeficientes caem abruptamente após *q lags*. Já o gráfico de PACF decai exponencialmente, conforme pode-se observar nas figuras 8 e 9.

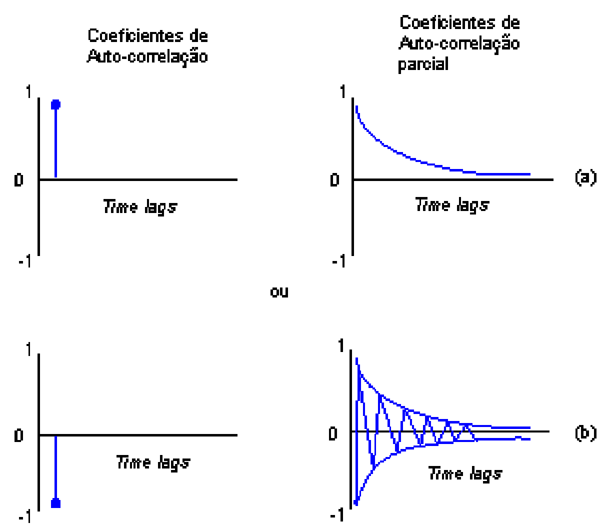


Figura 8 Gráficos Teóricos de ACF e PACF de uma série, indicando um modelo MA(1)
FONTE: Adaptado do livro Forecasting Methods and Applications, Makridakis 1978.

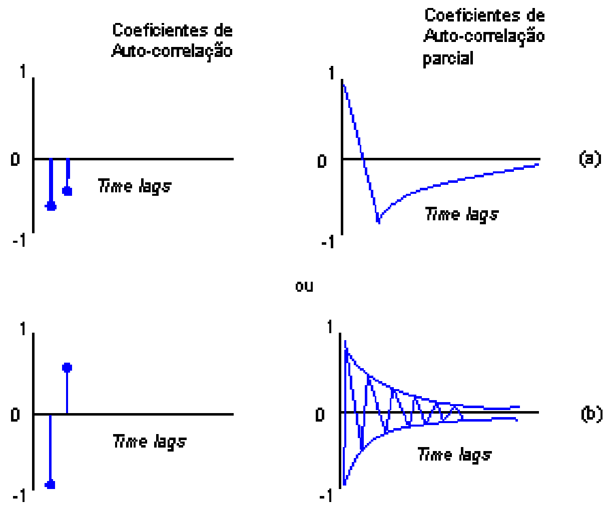


Figura 9 Gráficos Teóricos de ACF e PACF de uma série, indicando um modelo MA(2).
FONTE: Adaptado do livro Forecasting Methods and Applications, Makridakis 1978.



Os modelos mistos ARMA (p,q), combinam as características e as propriedades dos dois modelos que a compõem. Ou seja, apresentam comportamentos similares nas funções de ACF e PACF, gerando gráficos semelhantes, com várias possibilidades de padrões.

Makridakis destaca que a principal etapa da metodologia Box-Jenkins é identificar o modelo ARMA (p,q) mais apropriado através da análise das funções ACF e PACF, mas não de forma automática pois necessita do julgamento humano.

Esta análise é dividida em três etapas:

1. Verificar se a série é estacionária ou não. Caso não seja, proceder com os devidos ajustes;
2. Avaliar os gráficos de ACF e PACF para se identificar o modelo mais adequado;
3. Analisar as correlações de forma a se definir o grau de cada modelo.

### 2.3.8.5. Modelos Sazonais

O método de Box-Jenkins também pode ser adotado para as séries temporais que apresentem padrões de comportamento periódico, ou seja, que se repetem em

*S* períodos de tempo (sazonalidade). Sua identificação segue o mesmo método descrito para as séries não sazonais. O modelo geral é dado pela equação 55, apresentada a seguir:

$$\begin{aligned}&(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p)(1-\Phi_1 B^S-\Phi_2 B^{2S}-\ldots-\Phi_P B^{PS})X_t = \\ &(1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q)(1-\Theta_1 B-\Theta_2 B^2-\ldots-\Theta_{QS} B^{QS})\varepsilon_t\end{aligned} \qquad \text{eq(55)}$$

Onde:

$\Phi$ e $\Theta$ - Parâmetros do modelo sazonal;

$\phi$ e $\theta$ - Parâmetros do modelo não-sazonal;

S – Comprimento da sazonalidade (S = 12 para dados mensais);

$(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p)$ - Refere-se ao termo auto-regressivo não-sazonal de grau p com operador de defasagem B;

$(1-\Phi_1 B^S-\Phi_2 B^{2S}-\ldots-\Phi_P B^{PS})$ - Refere-se ao termo auto-regressivo sazonal de grau P com operador de defasagem B;

$(1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q)$ - Refere-se à parte não-sazonal de médias móveis de grau q com operador de defasagem B;

$(1-\Theta_1 B-\Theta_2 B^2-\ldots-\Theta_{QS} B^{QS})$ Refere-se à parte sazonal de médias móveis de grau Q com operador de defasagem B;

$\varepsilon_t$ - Erro.

Desta forma, tem-se a representação do modelo ARMA (p,q) (P,Q) onde *p* e *q* referem-se aos graus dos polinômios dos modelos auto-regressivo e de médias móveis não-sazonais, respectivamente. Já *P* e *Q* referem-se aos graus dos modelos auto-regressivo e médias móveis sazonais.

Cabe destacar que para se aplicar a metodologia de Box-Jenkins nas séries temporais com sazonalidade, também se faz necessário que a série seja estacionária, ou seja, tem que ter média e variância constantes ao longo da série e covariância como função da defasagem entre os *lags*.

No caso da série temporal não ser estacionária e possuir sazonalidade, teremos um modelo ARIMA (p,d,q) (P,D,Q), onde *d* representa a ordem de diferenciação não sazonal e *D* a ordem de diferenciação sazonal. Estas diferenciações, a exemplo do modelo sem sazonalidade, ocorrem da seguinte forma:

$$W_t = \nabla^d X_t = X_t - X_{t-1} \text{ (diferença)} \qquad \text{eq(56)}$$

$$W_t = \nabla^D X_t = X_t - X_{t-S} \text{ (diferença sazonal)} \qquad \text{eq(57)}$$

A definição da estacionaridade será feito observando-se a função de autocorrelação da série. Esta função também mostrará se há componente sazonal, o que é observado quando esta segue um padrão gráfico de picos periódicos.

## 2.4. MEDIDAS DE ERRO NA PREVISÃO

O processo de previsão da demanda tem como pressuposto básico que, ao se utilizar informações passadas, obtem-se uma boa diretriz sobre o que ocorrerá no futuro. Contudo, não existe uma garantia que o conhecimento sobre o passado trará informações exatas sobre este futuro. Sempre existirá uma incerteza vinculada às previsões.

Desta forma, não se pode simplesmente implementar um modelo de previsão de demanda sem acompanhar o seu desempenho ao longo do tempo.

Quando um modelo de previsão é selecionado, busca-se a maior precisão possível, ou seja, a menor diferença que possa existir entre a demanda realizada e a previsão. Esta diferença é chamada erro de previsão ou resíduo e da mesma forma que os dados passados trazem informações valiosas sobre o comportamento futuro, também deve-se considerar que os erros decorrentes das previsões devem ser avaliados com bastante critério, pois também podem trazer questões importantes para a gestão do processo em questão.

Dois aspectos justificam a análise destes erros:

1. Com base no comportamento dos erros, pode-se chegar à conclusão que é necessário o ajuste do modelo ou até mesmo a troca do método inicialmente selecionado. Em outras palavras, trata-se do "nível de desempenho" do método aplicado;
2. Por outro lado, uma estimativa precisa de erro pode auxiliar ao gestor a montar planos de contingência, como por exemplo, níveis de estoques de segurança ou políticas de compras.

Conforme conceituado anteriormente, a fórmula básica do erro é:

$$E_t = R_t - P_t \qquad \text{eq(58)}$$

Onde:

$P_t$ - Previsão para o período *t*;

$R_t$ - Demanda real ocorrido em *t*;

Obviamente, o erro negativo indica que a previsão gerada ficou acima do realizado, assim como o erro positivo mostra que a previsão ficou abaixo do realizado.

Esta definição de "erro" é, em geral, chamada em Estatística de "resíduo", pois corresponde a uma quantidade observável, enquanto o "erro" de um modelo é uma variável aleatória não observável, supostamente com média zero e variância constante.

Existem vários indicadores disponíveis para a mensuração da precisão dos métodos de previsão. A seguir alguns deles serão apresentados:

### 2.4.1. Erro Absoluto Médio (MAD)

O Erro Absoluto Médio, também conhecido como MAD (Mean Absolute Deviation) é um indicador bastante adotado, utilizado quando o gestor quer avaliar o erro na mesma unidade de medida da série original. O valor absoluto do erro é utilizado para se evitar que os erros positivos anulem os erros negativos. A sua fórmula é dada por:

$$MAD = \frac{\sum_{t=1}^{n} \left| R_t - P_t \right|}{n} \qquad \text{eq(59)}$$

### 2.4.2. Erro Percentual Absoluto Médio (MAPE)

O Erro Percentual Absoluto Médio , ou MAPE (Mean Absolute Percentual Error) é o erro médio em porcentagem, ao invés de quantidade. Este indicador é útil para determinar a amplitude do erro da previsão em relação aos valores da série histórica. Sua fórmula é dada por:

$$MAPE = \frac{\sum_{t=1}^{n} \left| R_t - P_t \right| / R_t}{n} \qquad \text{eq(60)}$$

### 2.4.3. Erro Quadrático Médio (MSE)

Outra medida de erro de previsão bastante utilizada é o Erro Quadrático Médio (EQM), em inglês, Mean Squared Error. Sua fórmula é dada por:

$$MSE = \frac{\sum_{t=1}^{n} (R_t - P_t)^2}{n} \qquad \text{eq(61)}$$

Esta medida amplifica o efeito dos *outliers,* que recebem uma grande importância quando deveriam ser desconsiderados da previsão. Outra característica deste indicador é que sua unidade de medida é a unidade do valor real ao quadrado. Este fato é resolvido adotando-se a raiz quadrada do MSE, ou RMSE (Root Mean Squared Error).

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (R_t - P_t)^2}{n}} = \sqrt{MSE} \qquad \text{eq(62)}$$

### 2.4.4. Viés Do Erro De Previsão

Denomina-se viés do erro de previsão quando um determinado modelo apresenta previsões consideravelmente acima ou abaixo do realizado. Em outras palavras, se o modelo subestima ou superestima a demanda. O ideal é que o viés da previsão oscile em torno de zero, o que demonstrará que os resíduos existentes no modelo são efetivamente aleatórios. Caso contrário, teremos um modelo que, muito provavelmente, está incorreto ou que seus parâmetros internos estejam errados.

A melhor maneira de se detectar o viés é através de montagem de um gráfico de soma acumulativa do erro de previsão (SC), que nada mais é do que a soma dos erros da previsão ( $E_t = R_t - P_t$) ao longo dos períodos.

A título de demonstração, foi gerada a previsão de demanda para o modelo de estojos para brinco argola grande (1001147). Adotou-se o método de Winters, contudo sem otimizar os coeficientes de amortecimento, o que gerará previsões distorcidas. Os coeficientes utilizados foram $\alpha = 0{,}1$, $\beta = 0{,}1$ e $\gamma = 0{,}1$.

Na tabela 2, a seguir, pode-se observar os resultados:

| ANO | MES | Realizado | Pré-visão | Erro | Erro acumulado | ANO | MES | Rea-lizado | Pré-visão | Erro | Erro acumulado |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2004 | 1 | 94 | 93 | (0,60) | (0,60) | 2006 | 1 | 100 | 93 | (6,76) | (62,82) |
| | 2 | 97 | 96 | (0,63) | (1,23) | | 2 | 73 | 69 | (4,10) | (66,92) |
| | 3 | 123 | 122 | (0,80) | (2,03) | | 3 | 106 | 101 | (4,81) | (71,73) |
| | 4 | 90 | 89 | (0,59) | (2,61) | | 4 | 84 | 81 | (2,95) | (74,69) |
| | 5 | 170 | 169 | (1,12) | (3,73) | | 5 | 139 | 135 | (3,55) | (78,24) |
| | 6 | 165 | 164 | (1,09) | (4,83) | | 6 | 155 | 152 | (2,58) | (80,81) |
| | 7 | 91 | 90 | (0,61) | (5,43) | | 7 | 87 | 86 | (0,73) | (81,55) |
| | 8 | 105 | 104 | (0,70) | (6,14) | | 8 | 146 | 146 | (0,13) | (81,67) |
| | 9 | 119 | 118 | (0,80) | (6,94) | | 9 | 136 | 137 | 0,81 | (80,87) |
| | 10 | 126 | 125 | (0,86) | (7,80) | | 10 | 96 | 97 | 1,15 | (79,72) |
| | 11 | 140 | 139 | (0,96) | (8,76) | | 11 | 88 | 90 | 1,53 | (78,19) |
| | 12 | 224 | 222 | (1,54) | (10,30) | | 12 | 213 | 218 | 4,69 | (73,50) |
| 2005 | 1 | 97 | 90 | (7,26) | (17,56) | 2007 | 1 | 96 | 92 | (4,27) | (77,76) |
| | 2 | 89 | 83 | (5,96) | (23,52) | | 2 | 80 | 77 | (2,68) | (80,44) |
| | 3 | 88 | 83 | (5,21) | (28,74) | | 3 | 131 | 128 | (3,05) | (83,49) |
| | 4 | 114 | 108 | (5,89) | (34,63) | | 4 | 152 | 150 | (2,09) | (85,58) |
| | 5 | 146 | 140 | (6,48) | (41,11) | | 5 | 200 | 199 | (1,01) | (86,59) |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 6 | 114 | 110 | (4,26) | (45,37) | | 6 | 234 | 235 | 0,68 | (85,91) |
| | 7 | 77 | 75 | (2,36) | (47,74) | | 7 | 136 | 137 | 1,36 | (84,55) |
| | 8 | 135 | 132 | (3,29) | (51,03) | | 8 | 241 | 245 | 3,95 | (80,60) |
| | 9 | 104 | 102 | (1,92) | (52,95) | | 9 | 185 | 189 | 4,06 | (76,54) |
| | 10 | 91 | 90 | (1,18) | (54,13) | | 10 | 147 | 151 | 3,93 | (72,60) |
| | 11 | 139 | 138 | (1,10) | (55,22) | | 11 | 174 | 179 | 5,36 | (67,24) |
| | 12 | 259 | 258 | (0,84) | (56,06) | | 12 | 347 | 359 | 11,86 | (55,38) |

Tabela 2 Medida de erro de previsão. Estojo para brinco argola grande

FONTE: Gerência da CAM

E para uma melhor visualização do conceito de viés, os dados referentes ao erro e erro acumulado apurados foram plotados no gráfico 10, a seguir.

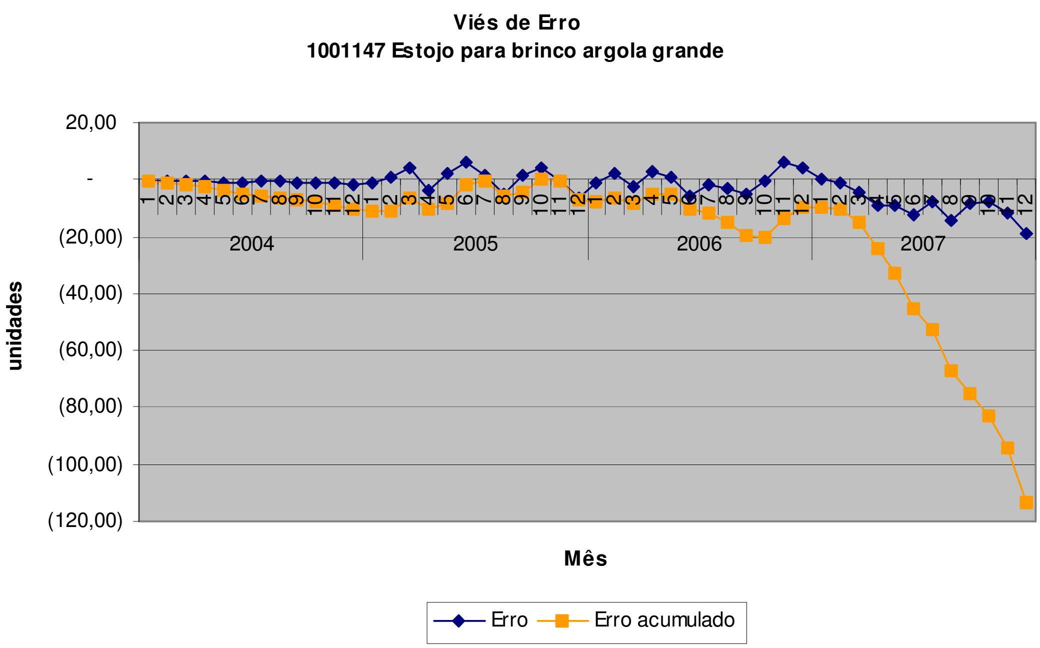


Figura 10 Viés de Erro de Previsão. Exemplo estojos para brinco argola grande.

FONTE: Gerência da CAM.

No exemplo indicado, o modelo necessita de ajuste, pois o erro acumulado mostra que as previsões estão abaixo da demanda real, ou seja, subestima a demanda com freqüência. Conforme mencionado anteriormente, o ideal é que o viés oscile em torno de zero. Provavelmente, neste caso o ajuste deverá estar relacionado diretamente com as constantes de amortecimento, visto que foram fixadas arbitrariamente em 0,1.

## 2.5.SOFTWARE

A rápida evolução da tecnologia observada nas últimas décadas, principalmente a adoção de microcomputadores, alavancou a utilização de sistemas computacionais para a análise e controle dos processos nas empresas. Diversas ferramentas têm sido criadas para auxiliar os gestores no processo de tomada de decisão.

O processo de previsão de demanda foi um dos que teve seu uso facilitado e disseminado em função deste processo de desenvolvimento tecnológico. Hoje existem no mercado diversos softwares que realizam cálculos estatísticos automaticamente, facilitando a vida dos analistas. Entretanto, a parte qualitativa da previsão ainda necessita do julgamento humano para atingir o seu objetivo. Em outras palavras, os softwares de previsão, em geral, funcionam como base para o processo previsão de demanda, fornecendo subsídios para a avaliação dos modelos e respectiva tomada de decisão.

Nas previsões deste estudo, foi utilizado o software *Forecast Pro*, adotado em grandes companhias nos EUA. Este software, a partir de uma série histórica fornecida, permite a identificação automática de modelos e também a rápida aplicação dos modelos univariados ajustados com o objetivo de previsão.

Os principais métodos matemáticos de previsão estão disponíveis no *Forecast Pro*, inclusive todos os detalhados no presente estudo, bem como alguns outros, como por exemplo, regressão dinâmica. Esta ferramenta possibilita rapidamente a avaliação de todos os indicadores de performance de cada método, facilitando sobremaneira o trabalho do analista.

## 2.6.
## PROCESSO

Segundo Makridakis (1978), ao contrário do que encontramos nas relações da Física, onde não existem erros ou variações aleatórias, nos sistemas econômicos ou comportamentais, a aleatoriedade está sempre presente. A melhor forma de se trabalhar, durante a modelagem, numa metodologia de previsão é tentar isolar esta aleatoriedade e minimizá-la. Busca-se, assim, o menor erro possível.

As séries históricas serão tratadas no *Forecast Pro* e, num primeiro momento, será permitido que o software indique o modelo mais adequado à série. A partir da análise dos indicadores de desempenho gerados pelo sistema, avalia-se a necessidade de alterações, que podem incluir desde a aplicação de eventos específicos, até a mudança do modelo selecionado previamente.

Foram selecionados dez modelos de estojos para análise, que representam 80% dos estojos consumidos anualmente na empresa.