2
Referencial Teórico

2.1.
Finanças Comportamentais

A seguir, há uma breve descrição de alguns conceitos e estudos referentes a Finanças Comportamentais, encontrados em alguns dos trabalhos já realizados (Lord et al. (1979), Kahneman e Tversky (1979), Rekenthaler (1998), Fama e French (1992), Bondt e Thaler (1985), Fama e French (1986), Cutler, Poterba e Summers (1991), Conrad e Kaul (1989), Fama (1991), Bernard e Thomas (1989), Michaely et al. (1995), French e Poterba (1991), Bernatzi e Thaler (2001), Barber e Odeam (2000), Odean (1998), Stein (1996), Baker e Wurgler (2002), Shefrin e Statman (1984), Thaler (1992), Lemgruber, Becker e Chaves (1988), Torres, Bonomo, e Fernandes (2001), Lucena, Andres e Ness (2003), Bonomo e Agnol (2003) e Lucena (2005)).

2.1.1.
Limitação à Arbitragem

A hipótese de mercado eficiente defende que os preços dos ativos refletem seus valores fundamentais, ou seja, que os preços refletem seus valores corretos. A base para essa definição é que todos os investidores são racionais, que possuem o mesmo número de informações e que têm expectativas homogêneas. Em um mercado eficiente, não há espaço para excesso de retorno, ou seja, retornos superiores aos riscos incorridos, ou até, no extremo, de obter retorno sem risco.

Estudos de Finanças Comportamentais defendem que os preços, na maioria das vezes, não refletem seus valores fundamentais, e que esse desvio justifica-se por existirem investidores que não são totalmente racionais. Essa definição é criticada pelas finanças modernas, que diz que os investidores racionais irão, rapidamente, desfazer qualquer descasamento, causado pelos investidores irracionais, entre o preço do ativo e seu valor fundamental. Sendo assim, de acordo com a teoria de mercado eficiente, caso o preço de um ativo esteja abaixo de seu preço justo, surge, então, a oportunidade para os investidores racionais investirem nesse ativo, com um risco extremamente baixo ou quase sem risco, pois a procura por esse ativo fará com que o preço suba até o seu valor fundamental. Esse é um caso conhecido como arbitragem, ou seja, obter retornos sem risco. Por isso, os investidores racionais são muitas vezes chamados de arbitradores. Os defensores comportamentais não concordam com essa teoria. Eles defendem que quando ocorre um problema de precificação de um investimento, as estratégias para corrigir esse preço podem ser arriscadas e de alto custo, permitindo, assim, que o preço permaneça abaixo ou acima do seu valor fundamental. Eles identificam três fatores que limitam a possibilidade de arbitragem, no caso em que os preços não estão eficientemente ajustados. O primeiro fator é o que eles identificam como risco fundamental, ou seja, a possibilidade do ativo em questão variar seu preço em função de novas informações. É o risco intrínseco ao ativo. O segundo fator é o risco de que os investidores não sejam totalmente racionais. Eles na maioria das vezes avaliam seus investimentos em função de suas performances. Em função de uma má performance, eles saem do mercado forçando uma piora nos preços. Além disso, muitos dos arbitradores são, também, agentes dos chamados investidores irracionais, que forçam que eles liquidem suas posições antes que o retorno se concretize. O terceiro ponto é a questão dos custos das operações. Além das taxas de corretagem e custódia, os arbitradores, muitas vezes, alugam ações e contraem financiamentos para realizarem suas operações. Como o tempo para recuperação de um ativo, por exemplo, pode ser mais longo ou bem mais longo do que o esperado, esses custos podem se tornar altos demais inviabilizando qualquer ganho de capital.

2.1.2. Psicologia: Sentimentos, Crenças e Preferências

De acordo com os defensores da teoria comportamental, o mercado é afetado pelos investidores irracionais como, por exemplo, no caso da limitação à arbitragem. Dessa maneira, alguns estudos, ligados à psicologia, foram feitos para entender melhor as formas de irracionalidade. Quais são os sentimentos, crenças e preferências das pessoas em geral que podem estar presentes nos investidores irracionais e que podem causar desvios no mercado.

Um dos sentimentos estudados foi o de excesso de confiança. Vários estudos concluíram que as pessoas apresentam excesso de confiança em seus julgamentos. Elas, normalmente, atribuem probabilidades maiores, do que as verdadeiras, de um certo resultado ou evento acontecer.

As pessoas, também, são, muitas vezes, otimistas e sonhadoras. A maioria delas apresenta uma visão acima da realidade com relação a suas habilidades e planos. Um estudo demonstrou que mais de 90% das pessoas pesquisadas pensam que estão acima da média em alguns assuntos como habilidade para dirigir, capacidade de se relacionar com outras pessoas e senso de humor.

Outra questão é a perseverança. De acordo com Lord et al. (1979), quando as pessoas formam suas opiniões, elas se apegam a essas opiniões e tendem a ficar com elas além do que deveriam, mesmo quando são apresentadas evidências que provem o contrário. Isso quer dizer que se as pessoas começarem a acreditar na eficiência do mercado, elas permanecerão acreditando, mesmo que apareçam evidências do contrário.

Normalmente, as pessoas formam suas estimativas a partir de um ponto inicial e, a partir dai, fazem os ajustes que julgarem pertinentes. Em um experimento foi questionada a quantidade de descendentes africanos nos Estados Unidos. Antes que os participantes respondessem, era perguntado se os percentuais que eles haviam estimados eram superiores ou inferiores a um certo percentual, definido de forma arbitrária. Os que foram perguntados para comparar suas estimativas com 10% apresentaram estimativas em torno de 25%. Já os que foram perguntados para comparar com 60% estimaram valores em torno de 45%. Esse fenômeno é conhecido como ponto de âncora.

Um ingrediente importante para entender os preços dos ativos é conhecer as preferências dos investidores e como eles reagem ao risco. A princípio os investidores deveriam seguir uma curva de utilidade (de acordo com a teoria da utilidade esperada), mas constantemente essa regra é quebrada no mercado acionário. Os estudiosos de finanças comportamentais acreditam que os casos em que não se observa a aplicação da teoria da utilidade esperada são de extrema importância para entender alguns dos fenômenos analisados por eles. Um desses casos foi batizado de teoria do prospecto [Kahneman e Tversky (1979)] ou aversão à perda [Rekenthaler (1998)]. Esses estudos evidenciaram que a questão da aversão à perda é mais relevante e decisiva que as decisões relacionadas ao risco. O homem irracional das finanças comportamentais, diante de uma perda certa, prefere se arriscar para tentar não perder, e diante de um ganho certo, prefere não se arriscar a um ganho maior. Outro aspecto relacionado com a aversão à perda é a questão do medo do arrependimento. Investidores irracionais tendem a não focar na alternativa de maior lucro por medo de terem que reportar perdas e reconhecerem seus erros, agindo, normalmente como a maioria do mercado. Esse comportamento pode levar, por exemplo, a duas situações. Uma é que o investidor hesita em vender ações por preços inferiores aos de compra, para evitar a concretização da temida perda. A outra é o conhecido efeito manada, onde o investidor copia o mercado, ou seja, se o mercado está comprando o papel A, o investidor compra também. Se o mercado está vendendo as ações de uma determinada empresa ou setor, o investidor, também, acompanha. Todas essas questões se contrapõem à teoria da utilidade, onde o investidor avalia o risco do investimento de acordo com a mudança que ele proporcionará em seu nível de riqueza.

2.1.3. Comportamentos do Mercado

Nesta seção serão apresentados alguns comportamentos encontrados no mercado de ações e que são difíceis de serem explicados através dos modelos racionais como, por exemplo, o CAPM (Capital Asset Pricing Model).

O primeiro comportamento é relativo aos retornos de ações de empresas menores quando comparados aos retornos de ações de empresas maiores. Fama e French (1992), dividiram as ações em dez portfólios com o mesmo número de papéis em cada um deles. No portfólio 1, foram agrupadas as menores empresas listadas na bolsa. Já no portfólio 10, foram colocadas as maiores empresas negociadas no mercado. Eles, então, acompanharam os retornos mensais desses portfólios, e verificaram que os retornos do portfólio 1 eram 0,74% por mês maior que os retornos do portfólio 10. Esse fato contrariava a teoria de risco retorno do CAPM, pois apesar dos betas das ações do portfólio 1 serem maiores que as do portfólio 10, essa diferença não era suficiente para explicar as diferenças entre os retornos.

Outro fato curioso é o denominado retorno ou reversão à média. Isso quer dizer que os retorno das ações oscilam em torno de um ponto médio. Sendo assim, os retornos dos papéis, de tempo em tempo, trocam de sinal. De maneira ainda mais clara e direta, papéis com retornos positivos por um certo período de tempo terão retornos negativos por um certo período e vice-versa. Esse fato pode ser verificado através de um estudo realizado por Bondt e Thaler (1985). Eles também criaram 2 portfólios de ações. Um portfólio com as 35 ações que tiveram maiores retornos em 3 anos, o chamado portfólio vencedor, e outro com as 35 ações com as piores performances em 3 anos, o chamado portfólio perdedor. Medindo os resultados desses portfólios nos 3 anos subseqüentes ao da formação dos portfólios, eles constataram que a média dos retornos das ações perdedoras era, aproximadamente, 8% por ano maior que os retornos dos papéis vencedores.

A possível previsibilidade dos retornos futuros de ativos também chama atenção. A teoria de mercados eficientes defende que os preços seguem um caminho aleatório (random walk), e que seria impossível, por exemplo, prever os retornos futuros das ações através de seus retornos passados. Entretanto, alguns estudos, como de Fama e French (1986), Cutler, Poterba e Summers (1991), Conrad e Kaul (1989) e outros mais, demonstram que há correlação entre a performance passada das ações e seus comportamentos futuros. Além dos retornos passados, conforme Fama (1991), outros fatores também foram relacionados à possibilidade de previsão de ativos como, por exemplo, relação preço-lucro, preço-valor contábil, anúncios de lucros ou dividendos, programas de recompra de ações e ofertas sazonais de ações por uma empresa. Bernard e Thomas (1989) fizeram um estudo relacionado ao anúncio de lucros pelas empresas, comparando os retornos das ações de empresas que anunciaram bons resultados, com os retornos das ações de empresas que divulgaram más notícias. As empresas que divulgaram bons resultados tiveram resultados, nos 60 dias seguintes, 4% acima do grupo das más notícias. Novamente, as diferenças nos betas dessas ações, não são suficientes para explicar a variação dos retornos. Michaely et al. (1995) chegaram a resultados parecidos com relação a anúncio de pagamento de dividendos. As empresas que anunciaram pagamento de dividendos tiveram resultados superiores àquelas que não o fizeram.

2.1.4.
Comportamento dos Investidores

A seguir, serão apresentados alguns dos comportamentos dos investidores irracionais identificados em diversos estudos na área de Finanças Comportamentais. Esses comportamentos não alteram, necessariamente, os preços dos ativos, mas são mais uma crítica ao modelo de mercado eficiente.

Um primeiro exemplo desses comportamentos é o fato de que os investidores não diversificam, suficientemente, seus investimentos, como recomendado por diversos modelos normativos. Os investidores, por exemplo, apresentam um viés doméstico, ou seja, preferem investir em ativos domésticos. Investidores nos Estados Unidos, Japão e Inglaterra alocam 94%, 98% e 82% de seus investimentos, respectivamente, em empresas nacionais, French e Poterba (1991). Uma explicação para isso pode ser o fato das pessoas preferirem situações familiares, onde elas se sentem em melhores condições de enfrentá-las do que as outras pessoas. Nesse caso, os investidores se acham mais familiarizados com o mercado doméstico do que com mercados de outros países.

Outra questão também relacionada à diversificação é a chamada diversificação ingênua. Muitos investidores costumam alocar 1/n de seus recursos nas n opções de investimentos disponíveis. Por exemplo, se existirem 2 fundos de renda variável e um fundo de renda fixa, o investidor irá alocar um terço de suas reservas em cada um destes fundos, se expondo excessivamente ao risco de renda variável, de acordo com Bernatzi e Thaler (2001).

De acordo com os modelos tradicionais, os volumes de negociações do mercado deveriam ser muito abaixo do que realmente são. O principal motivo para isso é que, em um mundo racional, se um investidor faz uma oferta de compra por uma determinada ação a um determinado preço e encontra alguém prontamente disposto a aceitar a proposta, o investidor deve desconfiar que o vendedor possui informações privilegiadas e não realizar o negócio. Além disso, o excesso de negociações faz com que haja um alto gasto com os custos transacionais, corroendo grande parte dos retornos. Segundo os estudiosos das finanças comportamentais o excesso de negócios está relacionado com o excesso de confiança dos investidores. Eles acreditam que possuem informações preciosas que justificam os negócios, mas que na verdade essas informações são fracas e não garantem nenhum sucesso nas operações. Consistente com essa teoria está a pesquisa realizada por Barber e Odeam (2000) que verificou que investidores com maior número de transações tiveram retornos bastante inferiores que a média dos retornos dos demais investidores.

Outro comportamento dos investidores identificado no mercado é o de manter a posse de ativos com resultados negativos por um tempo longo demais. Por outro lado, eles se desfazem rapidamente de ativos vencedores. Odean (1998) analisou 165 mil contas de clientes de uma corretora de ações, e encontrou que os investidores realizavam seus lucros numa proporção 68% maior do que realizavam perdas. Sendo assim, uma ação com retorno positivo tem 68% mais chances de ser vendida do que uma ação com desempenho negativo. Uma explicação irracional para isso é que os investidores acreditam, intuitivamente, na teoria de retorno à média.

A decisão de compra das ações, por sua vez, está igualmente dividida entre ações vencedoras e perdedoras. A escolha pela compra de um determinado papel está muito relacionada ao quanto essa determinada ação chamou atenção. A idéia é que uma pessoa não analisa todas as ações listadas para encontrar uma boa compra. Elas, provavelmente, compram ações que, de alguma maneira, chamaram suas atenções, sendo, muitas vezes, os motivos das atenções um excelente ou péssimo retorno passado. Essa idéia da atenção não se aplica às decisões de venda, porque, nesses casos, as escolhas estão limitadas às ações existentes do investidor em questão.

2.1.5. Finanças Comportamentais e as Organizações

Outra pesquisa em desenvolvimento na área comportamental trata de identificar como os investidores irracionais, tratados nos tópicos acima, afetam as finanças das organizações. De acordo com Stein (1996), por exemplo, os gestores racionais, sabendo da existência de investidores irracionais, devem emitir mais ações, caso acreditem que o preço da ação esteja muito alto, e recomprar ações, caso os papéis estejam desvalorizados. Esse comportamento é conhecido como momento de mercado. O momento de mercado pode explicar também a diferença de estrutura de capitais das empresas. Suponha duas empresas com, basicamente, as mesmas características, tamanho, lucratividade e ativos. Suponha, também, que a empresa A, em algum período no passado, teve a relação valor de mercado – valor de livro (market-to-book ratio) muito acima da empresa B. De acordo com o momento de mercado, a empresa A terá uma participação maior de capital próprio do que a empresa B. Essa teoria foi confirmada em um estudo publicado por Baker e Wurgler (2002).

Outra prática das empresas que pode derivar do comportamento dos investidores irracionais é o pagamento de dividendos pelas organizações. Como dividendos, historicamente nos Estados Unidos, são mais tributados que os ganhos de capital, os investidores, que pagam impostos, deveriam preferir que as corporações recomprassem suas ações, ao invés de receber dividendos por elas. Mas, na prática, os investidores parecem preferir o pagamento de dividendos. As Finanças Comportamentais tentam explicar esse fenômeno de algumas maneiras. Uma delas como sendo uma questão de autocontrole. Como muitas pessoas têm dificuldades em se policiar, uma das maneiras encontradas para aumentar o controle é de definir regras como, por exemplo, não gastar mais do que uma certa quantia por dia, só comer certos alimentos em determinados dias da semana ou só gastar os dividendos e não tocar no principal investido. De acordo com Shefrin e Statman (1984), outra explicação é que as firmas pagam dividendos para diminuir as chances de arrependimento dos investidores e, conseqüentemente, reduzir as chances dos investidores perderem o interesse nas ações da empresa. Suponha que a empresa não pague dividendos. O investidor, para realizar o ganho financeiro, deverá vender seus papéis. Se após a venda os papéis tivessem uma boa performance, os investidores ficariam bastante arrependidos. Para evitar esse tipo de sentimento os investidores procuram, então, investir em empresas que pagam dividendos.

Os dois parágrafos acima apresentam comportamentos dos gerentes ou das organizações em função da existência de investidores irracionais. Existem também gerentes irracionais que também causam impactos nas decisões de investimento das organizações e, obviamente, em seus resultados. Uma importante questão nesse contexto é o excesso de confiança ou otimismo dos gerentes. O impacto desses sentimentos pode ser verificado quando se analisa a compra de uma empresa por outra. As organizações, quando estão interessadas em adquirir uma outra empresa, fazem suas avaliações levando em conta os possíveis ganhos de sinergia, caso o negócio se concretize. Verifica-se na prática, porém, que os valores estimados dessas sinergias superam seus valores verdadeiros. Essa discrepância é explicada pelo excesso de otimismo e autoconfiança dos gestores. Dessa maneira, uma demonstração do excesso de otimismo por parte de algumas empresas pode ser verificada na análise de processos de venda tipo leilão, onde as diferenças entre as ofertas vencedoras e perdedoras podem chegar a valores exorbitantes, 143%, conforme Thaler (1992) constatou.

2.1.6. Trabalhos Realizados no Mercado Brasileiro

A seguir, serão apresentados os principais resultados de alguns trabalhos desenvolvidos sobre o mercado brasileiro de ações, com o intuito, não só, de se entender melhor o comportamento dos papéis, mas também para que se tenha uma idéia do avanço das pesquisas brasileiras sobre o tema. A apresentação será feita de maneira bastante breve e resumida. Sendo assim, caso o leitor queira um maior detalhamento sobre os testes e resultados apresentados, ou fazer uma investigação sobre itens omitidos em função da brevidade, sugere-se a consulta na integra dos trabalhos, cujas referências encontram-se mencionadas na seção referências bibliográficas.

Lemgruber, Becker e Chaves (1988) realizaram testes para verificar o efeito do final de semana no retorno diário das ações brasileiras, através da comparação dos retornos dos diferentes dias da semana. A amostra utilizada é referente aos índices IBV e IBOVESPA entre 17-ago-1983 e 24-ago-1987. Existem, basicamente, duas hipóteses para uma possível diferença entre os retornos de cada dia. A primeira de que os retornos das ações são gerados por "dias-calendário", ou seja, os retornos das ações teriam um comportamento ininterrupto e a expectativa de retorno para as segundas-feiras seria equivalente ao acumulado de três dias (sábado, domingo e segunda-feira). A outra hipótese de que os retornos das ações são gerados por "dias de negócio", ou seja, somente durante os dias úteis. Desta maneira, os retornos esperados para as segundas-feiras e os demais dias da semana seriam os mesmos. Eles verificaram, porém, que os retornos das segundas e terças feiras são, estatisticamente, nulos e os retornos para os demais dias da semana são significativamente positivos, não identificando, portanto, nenhuma das duas hipóteses. O comportamento encontrado por eles é conhecido como efeito de fim de semana.

Torres, Bonomo, e Fernandes (2001) encontraram em seus estudos evidências que contrariam o modelo de caminho aleatório, analisando horizontes diários e semanais das ações brasileiras, entre o período de 4-mar-1986 a 15-abr-1998. Além disso, eles verificaram, a partir das análises das auto-correlações cruzadas de primeira ordem entre os retornos de carteiras de firmas agrupadas segundo seu tamanho, que retornos de firmas grandes ajudam a prever retornos de firmas pequenas.

Lucena, Andres e Ness (2003) realizaram um trabalho para testar o pressuposto de normalidade e a questão da independência dos retornos diários das ações. Os testes estatísticos, Kolmogorov-Smirnov e Ljung-Box, foram aplicados sobre as ações brasileiras para o período entre 01/01/1999 a 17/10/2002, sendo rejeitada a hipótese de normalidade para 90% da amostra, considerando-se um intervalo de confiança igual a 95%, e verificada a existência de interdependência entre os retornos em, até, 80% dos casos.

Bonomo e Agnol (2003) lançaram-se em verificar se a adoção da estratégia conhecida como estratégia contrária, ou seja, comprar uma carteira de ações perdedoras, financiadas com a venda de uma carteira de ações ganhadoras, geram retornos anormais no mercado brasileiro. Eles adotaram, para realizar os testes, uma metodologia desenvolvida por Chopra et al (1992), com pequenas adaptações para adequá-la ao mercado acionário brasileiro. Eles constataram que, após 12 meses de formação das carteiras perdedoras e ganhadoras, as carteiras perdedoras tiveram um retorno de, cerca, de 19% acima do retorno das carteiras ganhadoras, sendo esse fato explicado pela reação exagerada dos investidores.

Lucena (2005) inicia seu trabalho testando os pressupostos de alguns modelos de finanças, como normalidade, utilizando o teste de Jarque-Bera, e estacionariedade, através dos testes de Dickey-Fuller e Philips-Peron. Em seguida, estudou, também, a relação do tamanho e "book-to-market ratio" nos retornos das carteiras de ações formadas através de quantis e análise de "clusters", em função da relação do tamanho e "book-to-market ratio", adotando o Modelo de Multifatores de Fama e French (1996). Posteriormente, fez pesquisas referentes às evidências de fenômenos, conhecidos na literatura, como “overreaction”, que diz respeito à reação exagerada dos preços dos ativos, ou melhor, dos investidores, à determinadas expectativas. Considerou, também, o “momentum”, que é conceituado como a tendência dos preços das ações ou ativos continuarem se movendo na mesma direção, após algum impulso; e abordou, também, o “downside risk”, que estuda a volatilidade dos retornos negativos passados. Para testar esses fenômenos, o autor utilizou o modelo, ligeiramente modificado, de Grinblatt e Moskowitz (2004). Por fim, ele testa o modelo modificado de Fama e French (1996), que é o modelo original acrescido da variância condicional dos retornos. Os resultados do trabalho são os seguintes: aproximadamente 80% das séries não apresentaram normalidade; quase todas as séries, com exceção de 3, são estacionarias; o autor verificou que, tanto pela separação por quantis, como por análise de clusters, o tamanho e o “book-to-market ratio” apresentaram ser relevantes no comportamento dos retornos das ações brasileiras; ele observou, também, a reação exagerada do mercado, em um momento eleitoral, e que o “downside risk” mostrou ser uma variável relevante, podendo ser utilizada para a demonstração da não eficiência fraca do mercado brasileiro; e, por último, constatou que, aparentemente, que o modelo modificado de Fama e French (1996) é adequado para previsão de retornos no mercado de capitais brasileiro.

Nesse capítulo serão apresentadas as principais teorias utilizadas neste trabalho. Primeiro, será abordada a questão do caminho aleatório. Em seguida, uma rápida abordagem do teste de normalidade de Jarque-Bera e o teste de estacionariedade, o teste aumentado de Dickey-Fuller. Para finalizar, um resumo da metodologia de Box&Jenkins, que inclui, além do teste de normalidade e estacionariedade, um estudo das funções de auto-correlação (FAC) e auto-correlação parcial (FACP), a definição do modelo a ser utilizado (AR, MA, ARMA ou ARIMA), as estimativas dos parâmetros dos modelos escolhidos, a validação do modelo e a utilização do modelo para a realização de estimativas.

2.2. Caminho Aleatório

Um dos modelos de caminho aleatório, de acordo com Souza (1996), segue a seguinte equação:

$$y_t = y_{t-1} + \varepsilon_t \qquad (\text{ou } \Delta y_t = \varepsilon_t).$$

Dessa maneira, y depende o seu valor passado mais um ruído branco ε_t, que tem média nula. A variável y percorre um caminho aleatório, quando os sucessivos valores de y são independentes e y se altera segundo alguma distribuição de probabilidade. De acordo com Souza (1996), o caminho aleatório pode ser apresentado, em termos estatísticos, da seguinte maneira:

$$P[y_t = y | y_{t-1}, y_{t-2}, ...] = P[y_t = y].$$

Uma definição bastante didática, intuitiva e informal é apresentada em Gujarati (2000):

> "O caminho aleatório é frequentemente comparado com o modo de andar de um bêbado. Ao sair do bar, o bêbado se desloca uma distância aleatória μ_t no tempo t, e se ele continuar a caminhar indefinidamente, no final vagueará cada vez mais longe do bar. Diz-se que ocorre o mesmo com os preços das ações. A cotação de hoje é igual à cotação de ontem mais um choque aleatório."

2.3. Teste de Normalidade

Um teste bastante utilizado para testar a normalidade de uma série é o chamado teste de Jarque-Bera (JB). Esse teste consiste na comparação da assimetria e curtose da série testada com a assimetria e curtose de uma normal, através da seguinte estatística:

$$JB = n\left[\frac{A^2}{6} + \frac{(C-3)^2}{24}\right]$$

Onde A é a assimetria e C a curtose.

Na verdade, o JB mede o excesso de assimetria e curtose com relação a normal. A estatística JB segue uma distribuição qui-quadrado com dois graus de liberdade. Para cada valor de JB temos uma probabilidade p associada à estatística qui-quadrado. Se o valor p for demasiadamente baixo, rejeitamos a hipótese nula de que a amostra é normal. Simplificadamente, pode-se assumir que o p valor é a probabilidade da amostra ser normal.

2.4. Teste da Estacionariedade

Estacionariedade significa que, em uma série, não existe crescimento ou declínio ao longo do tempo. Os valores devem estar distribuídos horizontalmente. Na prática, o que acontece é que os dados flutuam ao longo de uma média constante, independentemente do tempo.

Para verificar se uma série é estacionária ou não pode-se aplicar o teste da raiz unitária. Se for encontrada uma raiz unitária em uma série, rejeita-se a hipótese nula de que a amostra é estacionária. Suponha que os valores de y possam ser encontrados a partir da seguinte regressão:

$$y_t = \rho y_{t-1} + \varepsilon_t$$

O problema da raiz unitária é encontrado, quando o coeficiente ρ é igual a um. Uma série temporal que apresenta uma raiz unitária é conhecida como uma série temporal de caminho aleatório, ou seja, um exemplo de uma série temporal não estacionária.

Um conhecido teste de raiz unitária é conhecido como teste de Dickey-Fuller. Esse teste consiste em estimar o valor de ρ e dividi-lo por seu erro padrão, sendo esse valor conhecido como a estatística τ de Dickey-Fuller. A estatística τ é semelhante à estatística t de Student, porém os valores críticos foram tabulados por Dickey e Fuller com base em simulações de Monte Carlo (o título do teste é, então, uma homenagem a seus descobridores). Para definir se a série é estacionária ou não, deve-se comparar a estatística τ com os valores críticos de Dickey-Fuller. Se o valor absoluto da estatística τ for maior que os valores críticos absolutos de Dickey-Fuller, não rejeita-se a hipótese nula de que a série tem uma raiz unitária.

O teste utilizado neste trabalho é uma variação do teste de Dickey-Fuller. O teste apresentado acima se aplica apenas a séries com regressão de defasagem 1. Para corrigir essa limitação, o próprio Dickey e Fuller definiram uma parametrização, ou regressão, para contemplar correlações de ordem p, ou seja, a única diferença é a equação de regressão utilizada para definir o coeficiente ρ. Essa nova versão do teste é conhecida como teste aumentado de Dickey-Fuller, e foi o teste utilizado neste estudo.

2.5. Metodologia Box & Jenkins

A metodologia de Box&Jenkins (1994) nada mais é que um modelo de previsão de séries temporais. Ele é um modelo univariado, ou seja, um modelo que se baseia em uma única série temporal:

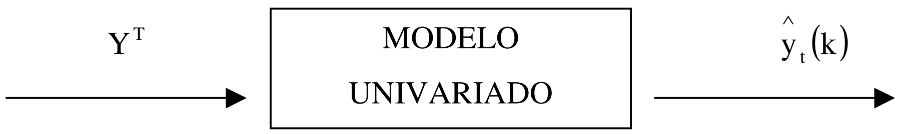

Figura 1 – Esquema ilustrativo de modelos univariados

Esta metodologia, como modelo univariado, tem a finalidade de encontrar uma equação que represente a série temporal y_t , por meio de uma estrutura dependente dos seus valores passados (processo auto-regressivo - AR) e seus erros de previsão um passo à frente (processo de média móvel - MA).

O exemplo mais simples de um processo auto-regressivo é o chamado AR(1):

$$y_t = \alpha_1 y_{t-1} + u_t$$

Onde u_t é um erro aleatório do tipo ruído branco. Sendo assim, o modelo diz que o valor de y no período t é uma proporção (α_1) de seu valor anterior no instante (t-1) mais uma perturbação ou erro aleatório. De forma análoga, o modelo AR(2) pode ser definido como segue:

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + u_t$$

Nesse caso, o valor de y depende de seus dois valores anteriores e um erro aleatório. Estendendo o conceito, a forma geral do modelo auto-regressivo, AR(p), ou seja, um processo auto-regressivo de p-ésima ordem, pode ser definida da seguinte maneira:

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + u_t$$

Por sua vez, o exemplo mais simples de um processo de média móvel é o denominado MA(1), ou processo de média móvel de primeira ordem:

$$y_t = \mu + u_t - \beta_1 u_{t-1}$$

Onde μ é uma constante e u é o termo de erro do tipo ruído branco. Esse processo define y no período t como sendo o valor de uma constante acrescida de uma média móvel dos termos de erros passado e corrente. O processo de média móvel de segunda ordem, MA(2), é definido pela equação:

$$y_t = \mu + u_t - \beta_1 u_{t-1} - \beta_2 u_{t-2}$$

E por fim, o caso geral de um processo MA(q) é definido como segue:

$$y_t = \mu + u_t - \beta_1 u_{t-1} - \beta_2 u_{t-2} - \dots - \beta_q u_{t-q}$$

Muitas vezes o termo y possui tanto características AR como características MA, ou seja, y é parcialmente explicado pelos seus valores passados, mas também pelos seus erros correntes e passados, são os conhecidos modelos ARMA. O modelo mais simples é o ARMA(1,1):

$$y_t = \theta + \alpha_1 y_{t-1} + u_t - \beta_1 u_{t-1}$$

Onde θ é uma constante. A forma geral é a ARMA(p,q), com p termos auto-regressivos e q termos média móvel:

$$y_t = \theta + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + u_t - \beta_1 u_{t-1} - \beta_2 u_{t-2} - \dots - \beta_q u_{t-q}$$

Os modelos AR, MA e ARMA são utilizados apenas para séries temporais estacionárias. Porém, caso uma série não seja estacionária, é possível diferenciá-la d vezes, até torná-la estacionária, e, então, aplicar o modelo ARMA(p,q). Esse é o chamado modelo ARIMA(p,d,q). ARIMA significa auto-regressivo integrado de média móvel.

A metodologia Box & Jenkins detalha os procedimentos necessários para a identificação do modelo, se AR, MA, ARMA ou ARIMA. Inicialmente é necessário rodar os testes de normalidade, pois para se aplicar a metodologia é necessário que a série siga uma distribuição normal. Em seguida, deve-se aplicar os testes de estacionariedade, para se verificar a necessidade de diferenciação, ou melhor, de transformar os dados em estacionários. A próxima fase é de identificação do modelo, ou seja, definir os valores de p e q. Posteriormente, vem a fase de estimar os parâmetros auto-regressivos e de média móvel, que, em seguida, devem ser testados na adequação do mesmo aos dados. O último passo é realizar as previsões de observações futuras. O resumo do fluxo da metodologia Box & Jenkins é apresentado na Figura 2, a seguir.

2.5.1. Identificação do Modelo

A identificação do modelo consiste em definir os parâmetros p, d e q do modelo ARIMA (p, d, q) apresentado acima. De acordo com Box et al. (1994), a definição dos parâmetros deve ser feita através das funções de auto-correlação (FAC) e de auto-correlação parcial (FACP) dos dados históricos.

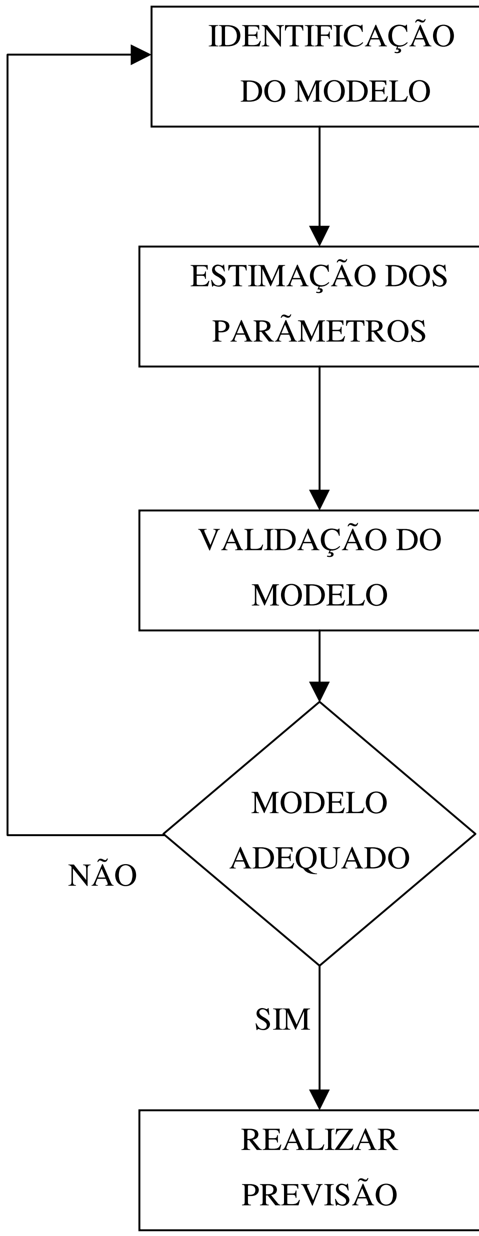

Figura 2 – Fluxograma da metodologia Box & Jenkins

A função auto-correlação deriva da função de auto-covariância definida por:

$$y_k = \mathbf{cov}(x_t, x_{t+k}) = E[(x_t - E[x_t]).(x_{t+k} - E[x_{t+k}])]$$

Como $E[x_t] = E[x_{t+k}] = \mu$ (média constante do processo estacionário):

$$y_k = E[(x_t - \mu).(x_{t+k} - \mu)]$$

A análise poderia ser feita a partir do comportamento das diversas auto-covariâncias, que revelaria a dependência entre os diversos termos da série. A função y_k, porém, não permite a comparação entre séries diferentes. Para contornar esse problema, foi definida, então, a função de auto-correlação:

$$\rho_k = \frac{y_k}{\mathbf{var}(x_t)} \qquad k = 0, 1, 2, ...$$

Sendo assim, ρ_k é uma medida padrão de dependência entre os intervalos de tempos (defasagens), onde, para todo k, $|\rho_k| \leq \mathbf{1}$.

A função auto-correlação parcial é definida como sendo o último termo auto-regressivo de um modelo AR(p):

$$\text{AR(1)} = y_t = \hat{\phi}_1 y_{t-1} + e_t$$

$$\text{AR(2)} = y_t = \phi_1 y_{t-1} + \hat{\phi}_2 y_{t-2} + e_t$$

$$\text{AR(3)} = y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \hat{\phi}_3 y_{t-3} + e_t$$

$$\text{AR(p)} = y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \ldots + \hat{\phi}_p y_{t-p} + e_t$$

Sendo assim, $\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \ldots, \hat{\phi}_p$ são os p coeficientes de auto-correlação parcial. Esses coeficientes podem ser calculados a partir das equações acima. Porém, essa tarefa seria muito trabalhosa e tomaria uma considerável quantidade de tempo. Uma maneira alternativa e mais satisfatória é calcular esses valores utilizando-se os próprios coeficientes de auto-correlação $\rho_{k's}$. Onde:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \phi_3 \rho_{k-3} + \ldots + \phi_p \rho_{k-p}$$

Considerando k=1, 2, 3,..., p e $\rho_i = \rho_{-i}$, se obtém um sistema de equações, conhecido como equações de Yule-Walker:

$$\begin{aligned}
\rho_1 &= \phi_1 + \phi_2\rho_1 + \dots + \phi_p\rho_{p-1} \\
\rho_2 &= \phi_1\rho_1 + \phi_2 + \dots + \phi_p\rho_{p-2} \\
&------------------- \\
\rho_p &= \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \dots + \phi_p
\end{aligned}$$

Portanto, conhecendo o valor de p e as auto-correlações $\rho_1, \rho_2, \dots, \rho_p$, pode-se calcular os coeficientes das auto-correlação parcial. Um método para a determinação da ordem p é aplicar sucessivamente as equações de Yule-Walker da seguinte forma:

$$\begin{pmatrix} 1 & \rho_1 & \dots\rho_{k-1} \\ \rho_1 & 1 & \dots\rho_{k-2} \\ ---- & ---- & ---- \\ \rho_{k-1} & \rho_{k-2} & \dots 1 \end{pmatrix} \text{x} \begin{pmatrix} \theta_{k1} \\ \theta_{k2} \\ ------- \\ \theta_{kk} \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ ------- \\ \rho_k \end{pmatrix}$$

Por fim, a auto-correlação parcial é a definida como a seqüência dos ϕ_{kk} obtidos com a resolução do sistema acima para k = 1, 2, 3, 4, ...

A identificação do modelo deriva, como já referido anteriormente, do comportamento da FAC e da FACP. A FAC decai exponencialmente para os modelos AR, sofre um corte brusco após a defasagem q e decai exponencialmente após a defasagem q para os modelos ARMA, ou seja, analisando a FAC é possível identificar apenas o parâmetro q. A definição do parâmetro p é feita através da análise da FACP, que, aliás, segundo Makridakis & Wheelwright (1978), só foi definida para essa finalidade. A FACP decai exponencialmente para os modelos MA, sofre um corte brusco após a defasagem p para os modelos AR e decai exponencialmente após a defasagem p para os modelos ARMA. A seguir um quadro (Figura 3) com o resumo das características das funções para os diversos processos:

	AR	MA	ARMA
Comportamento FAC	Decai exponencialmente	Corte brusco após a defasagem **q**	Decai exponencialmente após a defasagem **q**
Comportamento FACP	Corte brusco após a defasagem **p**	Decai exponencialmente	Decai exponencialmente após a defasagem **p**

Figura 3 – Comportamento da FAC e FACP nos modelos ARMA

Para que fique claro o entendimento da Figura 3, seguem-se, abaixo, exemplos dos gráficos das funções teóricas das FACs e das FACPs dos modelos AR(1), MA(1), ARMA(1,1) - (Figuras 4 a 9):

Modelo AR(1): exemplo 1

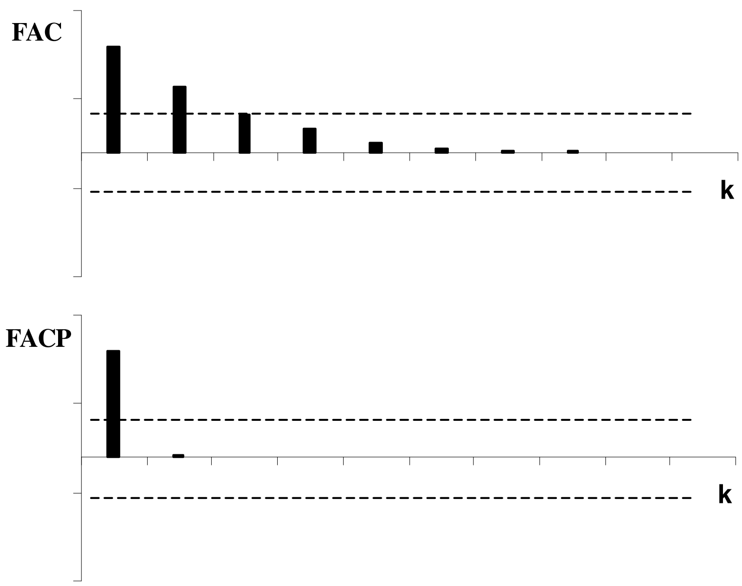

Figura 4: Exemplo 1 de FAC e FACP de um modelo AR(1)

Modelo AR(1): exemplo 2

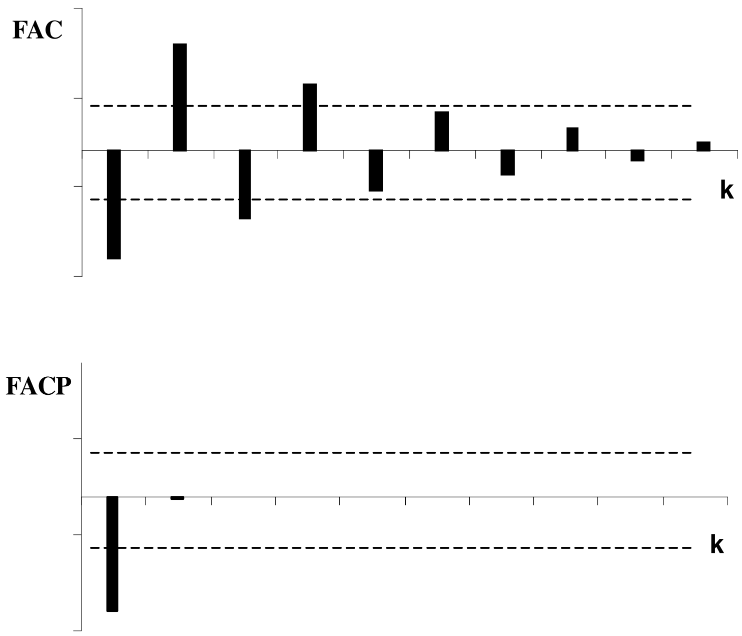

Figura 5: Exemplo 2 de FAC e FACP de um modelo AR(1)

Modelo MA(1): exemplo 1

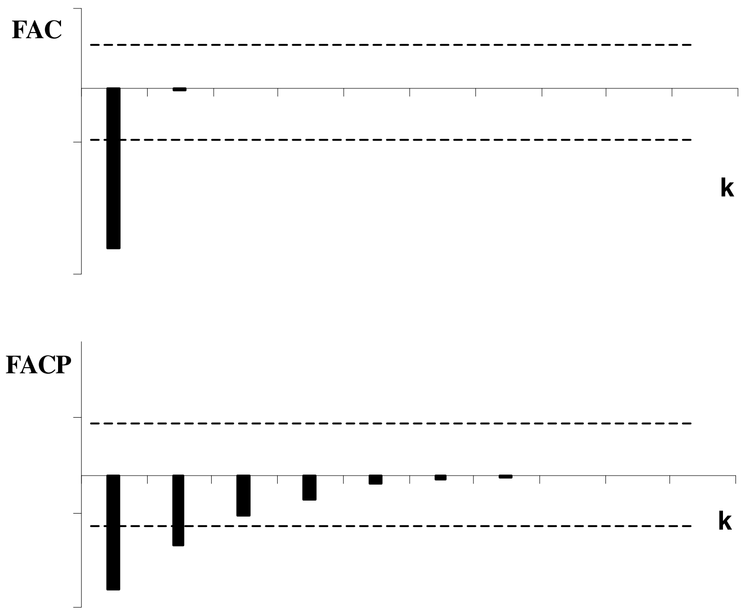

Figura 6: Exemplo 1 de FAC e FACP de um modelo MA(1)

Modelo MA(1): exemplo 2

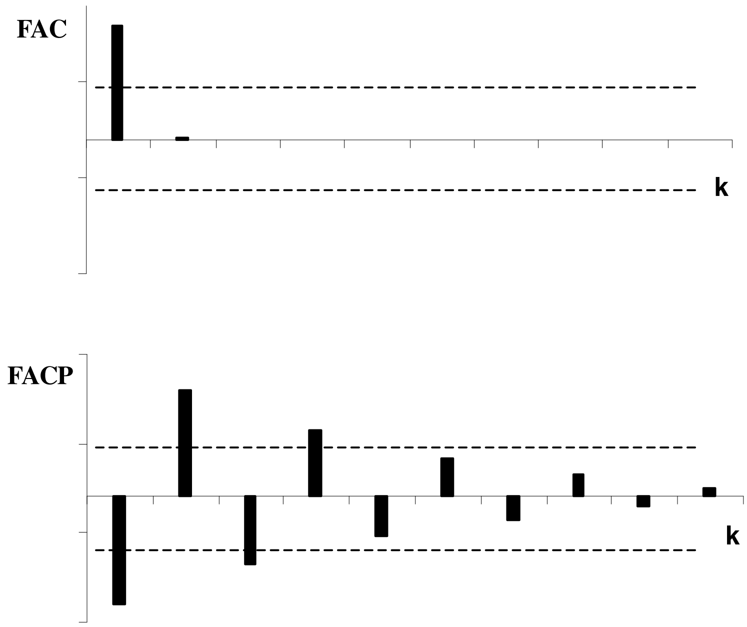

Figura 7: Exemplo 2 de FAC e FACP de um modelo MA(1)

Modelo ARMA (1,1): exemplo 1

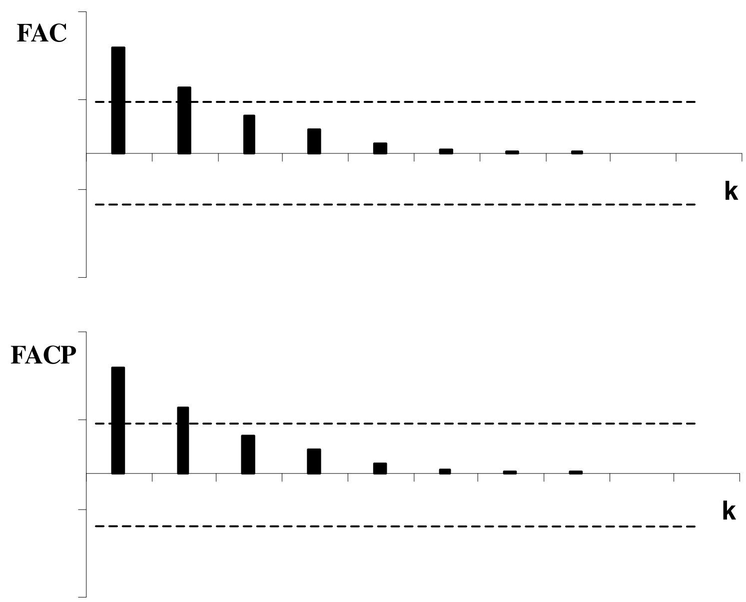

Figura 8: Exemplo 1 de FAC e FACP de um modelo ARMA (1,1)

Modelo ARMA (1,1): exemplo 2

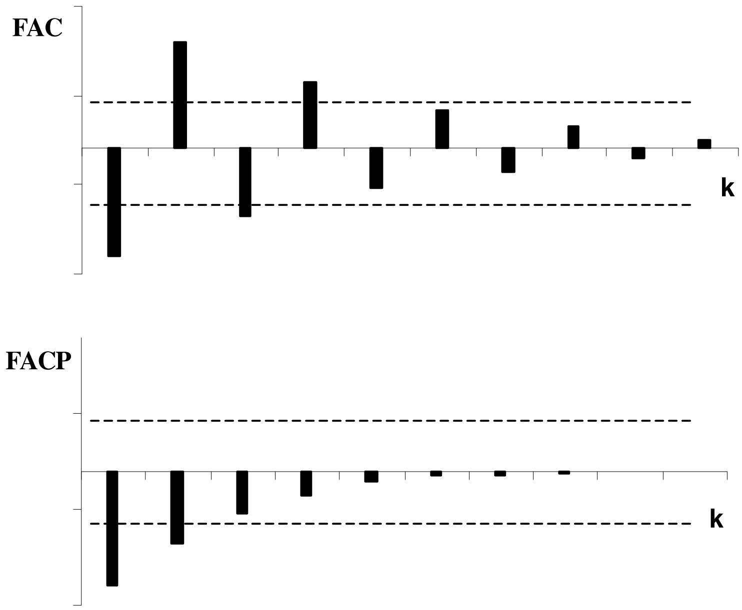

Figura 9: Exemplo 2 de FAC e FACP de um modelo ARMA (1,1)

Para finalizar a identificação, falta apenas a definição do parâmetro d. Na verdade, o processo de identificação do modelo ARIMA deve iniciar com o parâmetro d. Sua definição é muito simples, e pode ser feita sob duas óticas diferentes. A primeira, pela ótica da estacionariedade, onde o parâmetro d nada mais é do que o grau de diferenciação necessário para tornar os dados estacionários. A segunda ótica é pela análise do comportamento da FAC e da FACP, onde a série é diferenciada até o grau d, de forma que os comportamentos das funções mencionadas sejam semelhantes a um dos pares de comportamento apresentados na Figura 3. Uma vez definido o parâmetro d, devem ser traçadas as funções FAC e FACP da série diferenciada d vezes, para, então, serem definidos os parâmetros p e q do modelo ARIMA (p,d,q).

2.5.2. Estimação dos Parâmetros

Após definir o grau do modelo ARIMA (p, d, q), é necessário estimar os parâmetros α's e β's da equação de regressão. Os parâmetros auto-regressivos (α)

são lineares, permitindo o uso de métodos mais simples de estimação como, por exemplo, o de mínimos quadrados ordinários. Já os parâmetros de média móvel (β) não são lineares, tornando o processo de estimativa mais complexo. O método usualmente utilizado para a obtenção dos parâmetros de média móvel é o de máxima verossimilhança.

O método de mínimos quadrados, de forma bastante resumida, consiste na obtenção de parâmetros ($\alpha_1, \alpha_2, ..., \alpha_p$) e ($\beta_1, \beta_2, ..., \beta_q$) que minimize a soma das diferenças entre os pontos observados na amostra e os pontos estimados pela equação de regressão com os respectivos parâmetros ao quadrado, ou seja, que o somatório dos resíduos (a_t) seja mínimo:

$$S\left(\vec{\alpha}, \vec{\beta}\right) = \sum a_t^2 \Rightarrow \text{seja mínimo.}$$

O método de verossimilhança consiste, basicamente, em selecionar aqueles estimadores que maximizam a probabilidade de estimar valores, a partir da regressão, iguais aos valores realmente observados na amostra, ou simplificadamente, encontrar estimadores que gerem valores que mais se assemelhem aos valores da amostra. Para os leitores que desejam maiores informações sobre as definições das funções de máxima verossimilhança e suas propriedades de estimação, recomenda-se leitura de Bolfarine e Sandoval (2001).

2.5.3. Validação do Modelo

Conforme descrito nas etapas anteriores, o pesquisador deve examinar, visualmente, os gráficos da série temporal, da função auto-correlação e da função auto-correlação parcial, para definir o modelo. O exame do gráfico da série pode sugerir a existência de algumas imperfeições nos dados e ajuda a verificação da existência de estacionariedade ou não. A observação dos gráficos da FAC e da FACP, seja da própria série ou da série diferenciada, permite identificar os parâmetros p e q do modelo ARIMA (p, d, q). Como a identificação visual desses valores, em muitos casos, não é óbvia, pesquisadores diferentes podem escolher diferentes parâmetros p e q, e mesmo um único pesquisador pode encontrar diferentes parâmetros e que todos eles sejam bastante plausíveis. Dessa maneira, é preciso estabelecer critérios para verificar se o modelo desenhado é válido e critérios para escolher o melhor modelo dentre possíveis alternativas.

Uma idéia fundamental no modelo de Box&Jenkins é o princípio de parcimônia. O princípio da parcimônia significa que quanto menos parâmetros puderem ser adicionados ao modelo, melhor será a qualidade da regressão e, conseqüentemente, das estimativas. A idéia é que ao se aumentar a quantidade de parâmetros, se aumenta também a possibilidade de erros. Esse princípio vai de encontro ao critério do R^2, pois quando se aumenta o número de parâmetros, o valor de R^2 tende a se aproximar de um, ou seja, que a regressão é ótima. A explicação para isso é que apesar de se aumentar o valor de R^2, há uma redução do grau de liberdade, o que, de acordo com Box et al. (1994), prejudica a realização de previsões.

Um critério para avaliar a qualidade das estimativas, ou melhor, decidir qual modelo utilizar como, por exemplo, um modelo com mais parcimônia ou um modelo menos parcimonioso, é o conhecido AIC (Akaike Information Criteria), que pode ser calculado da seguinte maneira:

AIC = T ln(soma dos quadrados dos resíduos) + 2n

Onde n é igual ao número de parâmetros estimados (p, d, q, ..) e T é igual ao número total de observações. Para que a comparação entre alternativas seja feita corretamente, T deve ser mantido constante, bem como o período no qual as observações foram feitas. A melhor alternativa será aquela que apresentar o menor valor de AIC. Quando se aumenta o número de parâmetros, se tem duas possibilidades. A primeira é que a soma dos quadrados dos resíduos seja reduzida, pois com a adição do parâmetro, o modelo está sendo melhorado. Nesse caso, a redução do erro deve ser suficiente para compensar o aumento provocado pela adição do parâmetro. E a segunda alternativa é que haja uma piora no valor da soma dos quadrados dos resíduos, ficando óbvio que a adição do parâmetro não trouxe nenhum benefício para o modelo.

Outro teste que deve ser aplicado para verificar a qualidade do modelo é o teste de Portmanteau. Esse teste consiste em analisar as auto-correlações entre os resíduos encontrados com a utilização do modelo:

$$\hat{\rho}_i\left(\hat{a}\right);\ i = 0, 1, 2, \ldots, k$$

Box&Jenkins, então, sugere o seguinte teste estatístico (o teste de Portmanteau):

$$Q(k) = T\sum_{i=1}^{k}\hat{\rho}_i^{2}$$

Onde T é igual ao número total de observações. Para o teste acima, a hipótese nula é $H_0 = \rho_1 = \ldots = \rho_k = 0$, e a hipótese alternativa é $\rho_i \neq 0$, para qualquer i $\in$ {1, 2,..., k}. Q segue, aproximadamente, uma distribuição Qui-Quadrado (x^2) com k graus de liberdade.

O teste acima foi modificado por Ljung e Box (1978), para aumentar o poder do teste em amostras finitas:

$$Q(k) = T(T+2)\sum_{i=1}^{k}\frac{\hat{\rho}_i^{2}}{T-i}$$

O valor de k pode alterar a performance do teste. Não há um valor certo de k a ser utilizado, mas de acordo com Tsay (2002), estudos feitos através de simulações sugerem que a escolha de k = ln(T) proporciona um aumento no poder do teste. Na verdade, o que se está tentando demonstrar é que os ruídos gerados pelo modelo são ruídos brancos, ou seja, independentes e identicamente distribuídos.

Existem outros testes de diagnóstico dos modelos como, por exemplo, a análise das auto-correlações dos resíduos e o teste de periodograma acumulado. Os leitores interessados podem consultar em Box et al.(1994), capítulo 8.

2.5.4.
Realizar Previsão

É importante iniciar esta seção lembrando que o modelo definido com a aplicação das etapas anteriores, identificação, estimação e validação, obviamente, não é perfeito, pois os passos anteriores estão sujeitos a erros. Sendo assim, o sucesso do processo de previsão depende, fundamentalmente, das fases de identificação e estimação.

Com o modelo ARIMA, pretende-se prever os valores futuros da variável dependente, em função de seus valores passados e do valor presente. Existem diversas formas de explicitar os modelos de previsão. A maneira mais simples e intuitiva é através da equação de diferenças, que já foi apresentada, anteriormente, para alguns modelos. A seguir as equações de previsões de alguns dos modelos ARIMA mais utilizados:

Modelo ARIMA (1, 0, 0) ou AR(1):

$$y_{t+1} = \alpha_1 y_t + u_{t+1}$$

Modelo ARIMA (0, 0, 1) ou MA(1):

$$y_{t+1} = \mu + u_{t+1} - \beta_1 u_t$$

Modelo ARIMA (1, 0, 1) ou ARMA (1, 1):

$$y_{t+1} = \alpha_1 y_t + u_{t+1} - \beta_1 u_t + \delta$$

Modelo ARIMA (1, 1, 0) ou ARI (1,1):

$$y_{t+1} = (1 + \alpha_1) y_t - \alpha_1 y_{t-1} + \delta$$

Modelo ARIMA (0, 1, 1) ou IMA (1,1):

$$y_{t+1} = y_t - \beta_1 u_t + \delta$$

Modelo ARIMA (1, 1, 1):

$$y_{t+1} = (1 + \alpha_1) y_t - \alpha_1 y_{t-1} + u_{t+1} - \beta_1 u_t + \delta$$

Como os valores de u_{t+1}, nas equações de previsões, não são conhecidos no tempo t, eles devem ser substituídos por zero.

A maior preocupação com relação às previsões é a acurácia dos valores estimados. A análise da precisão das previsões também pode, e deve, ser utilizada para aprovar ou descartar um modelo. Existem diversas medidas de erro de previsão, não havendo um consenso de qual seria a melhor e que deva ser utilizada na prática.

Uma medida bastante utilizada é a representação dos desvios em termos relativos. A forma mais direta desse tipo de representação, de acordo com Makridakis e Wheelwright (1978), é apresentar os erros percentuais:

$$PE_t = \left[\frac{y_t - F_t}{y_t}\right] x\mathbf{100}$$

Onde F_t é o valor previsto de y no instante t.

Esta equação é utilizada para medir o erro percentual em qualquer instante t, mas não apresenta uma medida do erro ao longo do tempo. Se os erros de cada instante forem simplesmente somados, os erros negativos irão compensar os erros positivos, e o erro médio ou erro total seria demasiadamente baixo. A alternativa para essa questão, segundo Makridakis e Wheelwright (1978), é calcular a média dos valores absolutos dos erros percentuais (MAPE):

$$MAPE = \frac{\sum_{i=1}^{n} |PE_t|}{n}$$