



Universidade Estadual da Paraíba
Centro de Ciências e Tecnologia
Departamento Estatística

Alessandra Barbosa Santos

Análise de séries temporais: Um estudo de casos sobre cheques compensados.

Campina Grande
Dezembro de 2012

Alessandra Barbosa Santos

Análise de séries temporais: Um estudo de casos sobre cheques compensados.

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Orientador:

Tiago Almeida de Oliveira

Campina Grande
Dezembro de 2012

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL – UEPB

S237a Santos, Alessandra Barbosa .
Análise de séries temporais [manuscrito] : Um estudo de casos sobre cheques compensados. / Alessandra Barbosa Santos. – 2012.
50f.: il. color.

Trabalho de Conclusão de Curso (Graduação em Estatística) – Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2012.
“Orientação: Prof. Dr. Tiago Almeida de Oliveira, Departamento de Estatística”.

1. Series temporais. 2. Cheques. 3. Metodologia de Box e Jenkins. I. Título.

21. ed. CDD 332.76

Alessandra Barbosa Santos

Análise de séries temporais: Um estudo de casos sobre cheques compensados.


Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Aprovado em: 11 / 12 / 12

Banca Examinadora:



Prof. Dr. Tiago Almeida de Oliveira
Orientador



Prof. Dr. Gustavo Henrique Esteves
Universidade Estadual da Paraíba



Prof. Msc. Silvio Fernando Alves Xavier
Junior
Universidade Estadual da Paraíba

Dedicatória

À minha família.

Tudo Posso, Naquele que me fortalece.

Agradecimentos

Agradeço a Deus, por me manter firme nesta etapa da minha vida, me dando conforto e me mostrando que tudo é possível, se quisermos de verdade.

Ao professor Juarez, que através do amor que tem ao curso, indiretamente, nos encoraja a continuar, mesmo quando os obstáculos, nos parecem muito difíceis.

Aos professores Ricardo, Diana Maia e Ana Cristina pelo exemplo de profissionalismo e pela dedicação aos seus alunos.

À professora Ana Patrícia, que me deu embasamento na realização do TCC e em especial ao professor e orientador Tiago que transmitiu-me calma e segurança, sendo de fundamental importância no processo de conclusão deste trabalho.

Aos colegas de turma que de alguma forma contribuíram na minha graduação, em especial, agradeço ao meu colega Erasnilson, pela sua disposição e disponibilidade em multiplicar seu conhecimento, se mantendo sempre presente nas horas que precisei.

Aos meus familiares, em especial, minha irmã Ana Paula, que foi de grande importância, no período inicial de minha graduação e minha mãe, que sempre está ao meu lado.

Ao meu filho Cário Shermann e ao meu esposo José Cário, que dentro de suas limitações, compreenderam os momentos de ausência.

Resumo

Em um mercado financeiro competitivo, que cresce aceleradamente, com mudanças que vão, desde a substituição de moedas até a criação de novas e promissoras formas de negócio, a instituição “Banco”, que tem como principal fundamento, a arrecadação de lucros, busca se preparar antecipadamente, para enfrentar as oscilações de cada período. Com isso, este trabalho aborda uma análise estatística baseada na compensação de cheques, visando dar embasamento estatístico para que se possa prever informações futuras que auxiliem as instituições nas tomadas de decisões. Com a aplicação feita na série temporal observada (cheques compensados no estado da Paraíba), foi possível identificar o comportamento não estacionário, sazonalidade de 12 meses e tendência decrescente. Com a utilização da metodologia de Box e Jenkins, nos passos de identificação, estimação, diagnóstico e previsão, e com o auxílio do *software* R (R core team, 2011), foi sugerido, a identificação do modelo, com o uso do processo iterativo *auto.arima*, no entanto, não foi satisfatório. Optou-se então pela seleção do modelo de forma empírica, utilizando-se das funções de autocorrelação e autocorrelação parcial, que indicou possíveis modelos que bem representassem os dados, comparou-se através da representação gráfica do teste de Ljung-Box, a hipótese de ruído branco, e como o principal objetivo da análise é prever valores futuros, o modelo foi eleito com base na medida de precisão MAE, que apresentou um BIC moderado. Com isso o modelo selecionado foi um SARIMA multiplicativo de ordem $(3,1,0) \times (2,0,2)$, com uso de constante e apenas uma diferença simples.

Palavras-chaves: séries temporais, cheques, metodologia de Box e Jenkins.

Abstract

In a competitive financial market, which has grown rapidly, with changes ranging from the replacement of coins to the creation of new and promising forms of business, the institution “ Bank ”, whose main plea, raising profits, search prepare in advance to deal with fluctuations in each period. Consequently, this paper presents a statistical analysis based check clearing, aiming to give statistical foundation so that you can predict future information to assist institutions in making decisions. With the application made in the observed time series (cleared checks in the state of Paraíba), it was possible to identify non stationary the behavior, seasonality, and of 12 months downward trend. using the methodology of Box and Jenkins, in steps of identification, estimation, prediction and diagnosis and with the aid of the *software* R (R core team, 2011), suggested the identification of the model, using the interactive process *auto.arima*, however, was not satisfactory. opted it was then the selection of the model empirically, using autocorrelation functions and partial autocorrelation, which indicated possible models that represent well the data was compared by plotting the Ljung-Box test, the hypothesis noise white, and how the main goal of the analysis is to predict future values, the model was chosen based on precision measurement MAE, which showed a moderate BIC. with that model selected was a multiplicative SARIMA of order $(3,1,0)$ *times* $(2,0,2)$, using constant and only one simple difference.

Could not identify the stationary behavior, seasonality of 12 months

Key-words: time series, checks, Box and Jenkins methodology.

Sumário

Lista de Figuras

Lista de Tabelas

1	Introdução	p. 11
2	Desenvolvimento dos meios de pagamento através dos tempos	p. 12
2.1	Origem e evolução da moeda até a criação do cheque	p. 12
2.2	Tipos de moeda	p. 13
2.2.1	Escambo	p. 13
2.2.2	Moeda mercadoria	p. 13
2.2.3	Metal	p. 14
2.2.4	Moedas antigas	p. 14
2.2.5	Moeda papel	p. 15
2.2.6	Papel moeda	p. 15
2.2.7	Cheque	p. 16
3	Fundamentação Teórica sobre séries temporais	p. 19
3.1	Séries temporais	p. 19
3.1.1	Estacionariedade	p. 20
3.1.2	Transformações	p. 22
3.2	Modelos para séries temporais	p. 23
3.2.1	Processos estocásticos	p. 23

3.2.2	Tipos de modelos	p. 24
3.2.3	Função de autocorrelação(FAC)	p. 25
3.2.4	Modelos autorregressivos - AR(p)	p. 26
3.2.5	Função de autocorrelação parcial (FACP)	p. 26
3.2.6	Modelos de médias móveis - MA(q)	p. 27
3.2.7	Modelos autorregressivos e de médias móveis - ARMA (p,q)	p. 27
3.3	Modelos não estacionários	p. 28
3.3.1	Modelo autorregressivo integrado de médias móveis - ARIMA (p,d,q)	p. 28
3.3.2	Identificação do modelo	p. 30
3.3.3	Estimação do modelo	p. 32
3.3.4	Diagnóstico do modelo	p. 32
3.3.5	Previsão do modelo	p. 33
3.3.6	Medidas de precisão para previsão do modelo	p. 34
3.4	Modelos Sazonais	p. 35
3.4.1	Modelos SARIMA	p. 36
3.5	Metodologia de Box e Jenkins	p. 37
4	Aplicação	p. 40
5	Conclusão	p. 47
	Referências	p. 48

Lista de Figuras

1	Temperatura do ar, de dado local, durante 24 horas.	p. 20
2	Série não estacionária quanto ao nível e inclinação.	p. 22
3	Um processo estocástico interpretado como uma família variáveis aleatórias.	p. 24
4	Série não estacionária quanto ao nível.	p. 29
5	Cheques compensados na Paraíba no período de Janeiro 1997 a Dezembro de 2010.	p. 41
6	Teste Ljung-Box para validação do modelo.	p. 42
7	Gráficos para selecionar possíveis modelos adequados.	p. 42
8	Gráfico para validação do modelo.	p. 44
9	Previsão para quatorze meses com intervalo de confiança.	p. 45

Lista de Tabelas

1	Identificação dos modelos ARIMA.	p. 30
2	Modelos associados aos respectivos critérios de seleção.	p. 43
3	Estimativas e Erros dos parâmetros do modelo selecionado.	p. 44
4	Previsão para quatorze meses com intervalo de confiança.	p. 45
5	Valores observados e previstos para série cheques compensados e o erro associado.	p. 46

1 Introdução

Ao longo do tempo, o mercado econômico nacional passou por profundas mudanças, gerando um acelerado desgaste das diversas formas de moeda. A partir do momento em que as pequenas organizações crescem, criando e cedendo lugar, do até então, tímido cenário financeiro, ao amplo e promissor campo de negócios, surgem os bancos, no qual, o principal objetivo, na figura de seus representantes, é render lucros. Houve então a necessidade por parâmetros, que pudessem antecipar informações minimizadoras de possíveis transtornos futuros. Assim, os estudos estatísticos vieram como grandes aliados, nas etapas de planejamento destas instituições, com ênfase em séries temporais, que são de grande importância nos estudos que envolvem dados econômicos, nos diversos produtos de interesse das instituições, já que, na maioria dos casos, os dados são analisados no decorrer do tempo.

Sendo assim, esse trabalho é direcionado a análise do comportamento de cheques, no que se refere a sua compensação, sendo utilizado uma série de dados observados no estado da Paraíba no período de Janeiro de 1997 à Fevereiro 2012, retirados do site Banco Central do Brasil (2012).

Com aplicação de séries temporais e uso do *software* R para auxiliar nos cálculos dos valores estimados e na geração de gráficos, que facilitou no processo de identificação, utilizou-se a metodologia de Box e Jenkins, em busca do modelo que melhor explique os dados. Através das funções de autocorrelação e autocorrelação parcial observou-se possíveis adequações de modelos e posteriormente utilizar-se-á critérios com o propósito de adotar o melhor possível. Após concluídas as fases de ajuste do modelo, foram feitas previsões, e posterior validação cruzada, com intuito de se aproximar dos valores reais dentro das margens de erro. Constatando-se, assim, que as tomadas de decisões, pelos interessados na análise, sejam feitas com prudência e contetamento nos negócios a serem estabelecidos para o futuro.

2 Desenvolvimento dos meios de pagamento através dos tempos

2.1 Origem e evolução da moeda até a criação do cheque

Na antiguidade, em comunidades primitivas, o comportamento do homem dentro do ambiente em que vivia era bem diferente do que se vê nos dias atuais. Ele preocupava-se apenas com a sobrevivência, não objetivando nenhuma forma de vaidade, nem tão pouco conquistar riquezas. Em certo período desta época, iniciou-se, ainda timidamente, uma divisão do trabalho o qual variava de um lugar para outro, devido aos costumes de cada grupo. Essa divisão do trabalho modificou-se com o tempo. Parte dos bens e serviços obtidos domesticamente passaram a ser produzidos fora de casa ou da comunidade, por pessoas que se especializavam em determinadas profissões. Mais tarde, surgiram as fábricas e o trabalho passou a ser assalariado, dando início ao modo de produção capitalista. Com a demanda de mão de obra e do consumo, iniciou-se o processo de surgimento da moeda.

Moeda é um instrumento ou objeto aceito pela coletividade para intermediar as transações econômicas, para pagamento de bens ou serviços (VACONCELLOS; GARCIA, 2009).

Com o aquecimento comercial no Brasil e utilização maciça da moeda, observava-se uma crescente busca pela aquisição de bens e serviços, o que automaticamente se refletia em riquezas. Desta forma, houve a necessidade não só pela criação da moeda em si, mas também, por outras formas de pagamento, entre elas o “cheque”, que chegou ao mercado financeiro após uma longa evolução econômica.

2.2 Tipos de moeda

2.2.1 Escambo

Primeira forma de comércio que se tem conhecimento, Caracteriza-se pela simples troca de mercadoria por mercadoria, sem equivalência de valor. Esta elementar forma de comércio foi dominante no início da civilização, tendo-se estabelecido com os primeiros grupos humanos. Fritsch (1996), afirma que todo homem subsiste por meio da troca, tornando-se de certo modo comerciante e assim é que a própria sociedade se transforma naquilo que adequadamente se denomina sociedade comercial.

O escambo foi utilizado como forma geral de comércio, mas, apresentava alguns problemas no que se refere ao desenvolvimento das atividades econômicas de uma maneira geral e exigia uma dupla coincidência de desejos. Outro problema diz respeito à indivisibilidade dos objetos nas trocas diretas, uma vez que, se por exemplo, uma pessoa fabricasse canoas e precisasse comprar um quilo de feijão, a troca não seria adequada.

A primeira revolução agrícola foi modificando o sistema baseado no escambo. A produção passou a diversificar-se com a introdução de utensílios de trabalho. A divisão social do trabalho começa a se manifestar e os integrantes do grupo ganham funções específicas, gerando sensíveis mudanças na vida social. A atividade econômica tornou-se mais complexa, o número de bens e serviços exigidos para satisfação das necessidades do grupo aumentou, por consequência, a “dupla coincidência de desejos” torna-se mais difícil. Daí, alguns bens de aceitação são eleitos como intermediários de trocas, exercendo, portanto, função de moeda.

2.2.2 Moeda mercadoria

Segundo Freitas (2004), algumas mercadorias, pela sua utilidade, passaram a ser mais procuradas do que outras. Aceitas por todos, assumiram a função de moeda, circulando como elemento trocado por outros produtos e servindo para avaliar-lhes o valor. Eram as moedas-mercadorias.

Entre os bens usados como moeda está o gado, o sal e o bambu. Lopes e Rossetti (1992), comentam que a moeda pode ser conceituada como um intermediário de troca servindo como medida de valor e que tem aceitação geral, essa aceitação geral é um fenômeno essencialmente social. Além disso, como a moeda representa um poder de aquisição, desde o momento em que é recebida até o momento em que é dada em pagamento de outra

transação, ela também se caracteriza como uma reserva de valor.

Com o passar do tempo, as mercadorias se tornaram inconvenientes às transações comerciais, devido à oscilação de seu valor, pelo fato de não serem fracionáveis e por serem facilmente perecíveis, não permitindo o acúmulo de riquezas.

2.2.3 Metal

Os metais preciosos passaram a sobressair por terem uma aceitação mais geral e uma oferta mais limitada, o que lhes garantia um preço estável e alto. Além disso, não desgastavam-se, eram facilmente reconhecidos, divisíveis e leves. Segundo Fritsch (1996), o metal era, ao contrário de uma vaca, divisível para ser trocado por diversas porções de vários materiais, podendo ser cortado na quantidade certa para a troca. Todas as antigas civilizações compreenderam desde logo a importância da moeda e entenderam que os metais reuniam importantes características para serem utilizados como instrumentos monetários. Entretanto, havia o problema da pesagem, em cada transação, os metais preciosos deveriam ser pesados para se determinar seu valor. Esse problema foi resolvido com a cunhagem, uma vez que era impresso na moeda o seu valor.

2.2.4 Moedas antigas

Surgem, então, no século VII A.C., as primeiras moedas com características das atuais: são pequenas peças de metal com peso e valor definidos e com a impressão do cunho oficial, isto é, a marca de quem as emitiu e garante o seu valor.

Fritsch (1996) observa que, a operação de verificar a autenticidade ou quilate é ainda mais difícil e mais tediosa e, a menos que uma parte do metal seja fundida no cadinho ou crisol, utilizando dissolventes adequados, é extremamente incerta qualquer conclusão que se possa tirar. E, no entanto, antes de se instituir a moeda cunhada, as pessoas que não se submetessem a essa operação difícil e tediosa estavam expostas às fraudes e imposições mais penosas, pois em vez do peso em libra de prata pura ou de cobre puro, estavam sujeitas a receber pelas suas mercadorias uma composição adulterada dos materiais mais ordinários e baratos, os quais, porém, em sua aparência se assemelhavam à prata ou ao cobre.

2.2.5 Moeda papel

O desenvolvimento de sistemas monetários demandou o surgimento de um novo tipo de moeda: a moeda-papel (certificados emitidos mediante depósito em metal precioso). Ela veio para contornar os inconvenientes da moeda metálica (peso, risco de roubo), embora valessem com lastro nela. Assim surgem os certificados de depósito, emitidos por casas de custódia em troca do metal precioso nela depositado. Por ser lastreada, essa moeda representativa poderia ser convertida em metal precioso a qualquer momento, e sem aviso prévio, nas casas de custódia (LOPES; ROSSETTI, 1992).

A moeda de papel evoluiu quanto à técnica utilizada na sua impressão. Hoje a confecção de cédulas utiliza papel especialmente preparado e diversos processos de impressão que se complementam, dando ao produto final grande margem de segurança e condições de durabilidade. De acordo com Lopes e Rossetti (1992), a passagem da moeda papel para o papel moeda é tida como uma das mais importantes e revolucionárias etapas da evolução histórica da moeda.

2.2.6 Papel moeda

A evolução dos instrumentos monetários não pararia com a descoberta da operacionalidade da moeda papel. Os certificados emitidos, devido à sua aceitação já generalizada, passaram a circular mais que as próprias peças metálicas. Seu valor não decorreria ainda da regulamentação oficial de sua emissão, mas simplesmente da confiança geral em sua plena conversibilidade. Essas emissões monetárias trariam vantagens para produtores, comerciantes e banqueiros. Os primeiros passaram a ter acesso a uma nova fonte de financiamento; os comerciantes obtinham créditos suficientes para a expansão de seus negócios e os banqueiros beneficiavam-se das receitas correspondentes aos juros. Evidentemente, essa passagem histórica da moeda papel para as primeiras formas de papel moeda ou de moeda fiduciária (notas bancárias emitidas a partir de operações de crédito, sem lastro metálico) envolveria consideráveis margens de risco. Esses riscos evidenciados conduziram os poderes públicos a regulamentar o poder de emissão de notas bancárias. O direito à emissão de notas, em cada país, seria confiado a uma única instituição bancária oficial, surgindo, assim, os Bancos Centrais (FREITAS, 2004). Daí por diante, o papel moeda passou a receber a garantia das disposições legais que envolviam a sua emissão, o seu curso e o seu poder liberatório. Sua aceitação geral como meio de pagamento passou a substituir as garantias metálicas que apoiavam a moeda papel.

2.2.7 Cheque

Com a supressão da conversibilidade das cédulas e moedas em metal precioso, o dinheiro cada vez mais se desmaterializa, assumindo formas abstratas. Uma destas formas é o cheque, que, pela simplicidade de seu uso e pela segurança que oferece, passou, progressivamente, a ser adotado por número sempre maior de pessoas nas atividades de seu dia a dia.

Foi na Inglaterra que ocorreu a fase embrionária da letra de câmbio e a partir do século XVII, o cheque tomou impulso se aprimorando e tomando o contorno do título que hoje representa. Foi criado em 1694, o Banco da Inglaterra, que só passou a ser banco de emissão em 1742. A partir dessa data, em virtude do Banco não poder mais emitir ordens de pagamento aos seus depositantes contra os depósitos dos mesmos, foi difundida a prática de serem sacadas contra os bancos, letras de câmbio à vista. O sistema inglês, define o cheque, como uma letra de câmbio à vista, sacada contra um banqueiro. Porém, a Lei de 14 de junho de 1865, que regulamentou o cheque na França, definiu-o como “o escrito que, sob a forma de um mandato de pagamento, serve ao sacador para efetuar a retirada, em seu proveito ou em proveito de um terceiro, de todos ou parte dos fundos disponíveis, levados a crédito de sua conta pelo sacado” (MARTINS, 2000).

Segundo Mamede (2003), o cheque é uma ordem de pagamento emanada de uma pessoa (emitente ou sacador) que mantém contrato com uma instituição bancária (sacado) para que esta pague, imediatamente (à vista), determinada importância ao beneficiário nomeado, a sua ordem ou, não havendo nomeação do beneficiário ou nomeando-se genericamente o portador aquele a apresentar. O beneficiário do cheque pode ser igualmente chamado de tomador.

Na legislação brasileira, a palavra cheque surgiu com a Lei nº 149-B, de 20 de julho de 1893, tendo referido título sido contemplado com a Lei especial nº 2.591, de 7 de agosto de 1912, que regulou sua emissão e circulação (FREITAS, 2004). Ocorre que em 1931 promulgou-se a Lei Uniforme, em Genebra, lei essa que regulava o cheque, sua emissão, as modalidades de cheque existentes, os prazos e particularidades para apresentação para pagamentos junto às instituições financeiras, as possibilidades de transferência dos direitos sobre o cheque. Esta lei está em vigor no Brasil e foi incorporada pela Lei nº 7.357, de 2 de setembro de 1985 (OLIVEIRA, 1999). A Lei n.º7.357/85 não estabelece um conceito propriamente dito para o cheque. No entanto, tal normativo estabelece requisitos essenciais que o caracterizam e o distinguem, conforme artigo 1º.

- Artigo 1º - O cheque contém:

I. A denominação “cheque” inscrita no contexto do título e expressa na língua em que este é redigido;

II. A ordem incondicional de pagar quantia determinada;

III. O nome do banco ou da instituição financeira que deve pagar (sacado);

IV. A indicação do lugar de pagamento;

V. A indicação da data e do lugar de emissão;

VI. A assinatura do emitente (sacador), ou de seu mandatário com poderes especiais.

“Parágrafo único”: A assinatura do emitente ou a de seu mandatário com poderes especiais pode ser constituída, na forma de legislação específica, por chancela mecânica ou processo equivalente.

Com base na referida lei, pode-se extrair a conceituação das pessoas envolvidas no referido título de crédito:

- (i) sacador ou emitente: é a pessoa que emite o cheque, é quem dá a ordem;
- (ii) sacado: é o banco ou instituição assemelhada para quem parte a ordem;
- (iii) tomador ou beneficiário, que pode ser em alguns casos simplesmente portador: é a pessoa a favor de quem é dada a ordem.

Esse documento, pelo qual se ordena o pagamento de certa quantia ao seu portador ou à pessoa nele citada, visa, primordialmente, à movimentação dos depósitos bancários. O importante papel que esse meio de pagamento ocupa, hoje, na economia, deve-se às inúmeras vantagens que proporciona, agilizando a movimentação de grandes somas, impedindo o entesouramento do dinheiro em espécie e diminuindo a necessidade de troco por ser um papel preenchido à mão com a quantia de que se quer dispor. Segundo Teixeira (1988) o cheque é classificado como a forma mais clássica de retirada de fundos a qual o cliente dispõe. No entanto para alguns especialistas de economia, tal conceito é ultrapassado, pois aos poucos o cheque tornou-se sinônimo de insegurança e burocracia, e nos dias atuais cartões magnéticos, transferências via *internet banking* e outras formas disponibilizadas pela tecnologia superam e muito a utilização da tradicional folha de cheque.

Atualmente distintos instrumentos exercem a função de moeda: as moedas metálicas, o papel moeda e as moedas de plástico (cartões de débito) são empregados para a realização de pagamentos de pequenos valores, ao passo que as transferências de valores mais elevados se dão pela via de compensação bancária das contas de depósitos através da moeda escritural (cheque) ou da moeda eletrônica - transferências diretas, realizadas por interfaces baseadas nas novas tecnologias de informação (transferências eletrônicas por meio de *internet*, terminais de auto-atendimento, etc.) (PAIVA; CUNHA, 2008).

O dinheiro, seja em que forma se apresente, não vale por si, mas pelas mercadorias e serviços que pode comprar. É uma espécie de título que dá a seu portador a faculdade de se considerar credor da sociedade e de usufruir, através do poder de compra, de todas as conquistas do homem moderno. A moeda não foi, pois, genialmente inventada, mas surgiu de uma necessidade e sua evolução reflete, a cada momento, a vontade do homem de adequar seu instrumento monetário.

3 Fundamentação Teórica sobre séries temporais

3.1 Séries temporais

Segundo Fischer (1982), uma série temporal é um conjunto de observações de uma variável dispostas sequencialmente no tempo. Conforme o conjunto obtido, pode-se classificar a série em contínua ou discreta. Diz-se que a série cronológica é discreta quando o conjunto de observações no tempo for finito ou infinito enumerável. Caso contrário, isto é, se o conjunto gerado for infinito não enumerável, diz-se que a série é contínua.

A grande maioria dos estudos econômicos que se valem do tratamento de séries temporais, utilizam séries discretas, onde as observações são geradas em um intervalo de tempo com amplitude constante.

Alguns exemplos de séries temporais são citados por Morettin e Tolo (2006):

- (i) valores diários de poluição na cidade de São Paulo;
- (ii) valores mensais de temperatura na cidade de Cananéia-SP;
- (iii) índices diários da Bolsa de Valores de São Paulo;
- (iv) precipitação atmosférica anual na cidade de Fortaleza;
- (v) número médio anual de manchas solares;
- (vi) registro de marés no porto de Santos.

Conforme a classificação das séries citadas, pode-se dizer que nos itens (i) à (v) tem-se séries temporais discretas, enquanto (vi) observa-se o exemplo de uma série contínua.

Na realidade, o que é chamado de série temporal é uma trajetória, dentre muitas que poderiam ter sido observadas. Será possível uma melhor compreensão através de uma visualização da Figura 1, dada por Morettin e Tolo (2006).

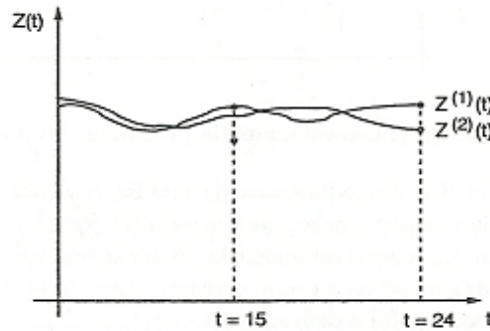


Figura 1: Temperatura do ar, de dado local, durante 24 horas.

Designando-se por $Z(t)$ a temperatura no instante t (dado em horas por exemplo), nota-se que para os dois dias diferentes, obtém-se duas curvas que não são, em geral, as mesmas. Estas curvas são chamadas trajetórias do processo físico que está sendo observado e este (o processo estocástico) nada mais é do que o conjunto de todas as possíveis trajetórias que podem ser observadas. Cada trajetória é também chamada uma série temporal ou função amostral.

3.1.1 Estacionariedade

Uma das suposições mais frequentes que se faz a respeito de uma série temporal é a de que ela é estacionária, ou seja, ela se desenvolve aleatoriamente, no tempo, em torno de uma média constante, refletindo alguma forma de equilíbrio estável (MORETTIN; TOLOI, 2006). Basicamente isso significa que o comportamento da série não se altera com o passar do tempo, e sua média não muda ao longo do tempo decorrido.

Segundo Ehlers (2007), o processo estacionário é uma importante classe dos processos estocásticos podendo ser estritamente estacionário ou fracamente estacionário.

- Série estritamente estacionária

Ocorre quando a distribuição de probabilidade conjunta de $Z(t_1), \dots, Z(t_k)$ é a mesma de $Z(t_1 + \tau), \dots, Z(t_k + \tau)$, ou seja, o deslocamento da origem dos tempos por uma quantidade τ não tem efeito na distribuição conjunta que portanto depende apenas dos intervalos entre t_1, \dots, t_k .

Em particular, para $k = 1$ a estacionariedade estrita implica que a distribuição de $Z(t)$ é a mesma para todo t de modo que, se os dois primeiros momentos forem finitos, tem-se que

$$\mu(t) = \mu \text{ e } \sigma^2(t) = \sigma^2$$

Para $k = 2$ a distribuição conjunta de $Z(t_1)$ e $Z(t_2)$ depende apenas da distância $t_2 - t_1$, chamada defasagem. A função de autocovariância $\Upsilon(t_1, t_2)$ também depende apenas de $t_2 - t_1$ e pode ser escrita como $\Upsilon(\tau)$. Onde,

$$\Upsilon(\tau) = E[Z(t) - \mu][Z(t + \tau) - \mu] = \text{cov}[Z(t), Z(t + \tau)]$$

é chamado de coeficiente de autocovariância na defasagem τ .

Verifica-se que o tamanho de $\Upsilon(\tau)$ depende da escala em que $Z(t)$ é medida. Portanto, para efeito de interpretação, é mais útil padronizar a função de autocovariância dando origem a uma função de autocorrelação

$$\rho(\tau) = \Upsilon(\tau)/\Upsilon(0),$$

que mede a correlação entre $Z(t)$ e $Z(t + \tau)$. Na prática é muito difícil usar a definição de estacionariedade estrita e costuma-se definir estacionariedade de uma forma menos restrita (EHLERS, 2007).

- Estacionariedade fraca ou segunda ordem

Ocorre quando sua função média é constante e sua função de autocovariância depende apenas da defasagem, i.e.

$$E[Z(t)] = \mu \text{ e } \text{Cov}[Z(t), Z(t + \tau)] = \gamma(\tau)$$

Nenhuma outra suposição é feita a respeito dos momentos de ordem mais alta. Além disso, fazendo $\tau = 0$, segue que $\text{Var}[Z(t)] = \gamma(0)$, ou seja a variância do processo, assim como a média também é constante. Observa-se também que tanto a média quanto a variância precisam ser finitas.

Estacionariedade fraca implica em estacionariedade estrita para processos normais. Por outro lado, μ e $\gamma(\tau)$ podem não descrever adequadamente processos que sejam muito “não-normais”.

- Séries não estacionárias

Em geral, é mais comum a existência de séries com comportamento inconstante, podendo crescer ou decrescer conforme o tempo passa, especialmente no setor de economia, onde é pouco provável encontrar uma série originada no tempo sob forma linear. “Um simples exame de algumas das variáveis mais comuns em Economia, como por exemplo, preço, produto nacional, renda, vendas, consumo, investimento e tantas outras, mostra que suas variações apresentam um comportamento não estacionário evolutivo no decorrer do tempo” (FISCHER, 1982).

O tipo de não estacionariedade mostrado por essas séries pode ser caracterizado como homogêneo. Ainda que as séries se movimentem livremente sem se fixarem em torno de um particular valor médio, seu comportamento em diferentes períodos de tempo é essencialmente similar, conforme ilustra a Figura 2, dada por Fischer (1982). Essas séries são conhecidas como séries não estacionárias.

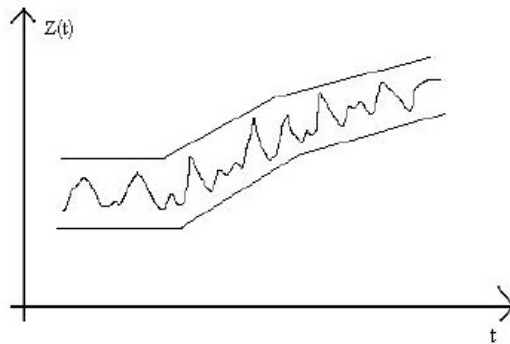


Figura 2: Série não estacionária quanto ao nível e inclinação.

Para efeito de solução, séries não estacionárias homogêneas apresentam a característica de terem propriedades tais que, diferenciando-as uma ou mais vezes, resultam em séries estacionárias. Será necessário recorrer a transformação desta série objetivando-se deixá-la estacionária.

3.1.2 Transformações

Segundo Morettin e Tolo (2006), como a maioria dos procedimentos de análise estatística de séries temporais supõem que estas sejam estacionárias, será necessário transformar os dados originais, se estes não formam uma série estacionária. Há basicamente, duas razões para se transformar os dados originais: estabilizar a variância e tornar o efeito sazonal aditivo.

A transformação mais comum consiste em tomar diferenças sucessivas da série original, até se obter uma série estacionária. A primeira diferença de Z_t é definida por

$$\Delta Z_{(t)} = Z_{(t)} - Z_{(t-1)}, \quad (3.1)$$

a segunda diferença é

$$\Delta^2 Z_{(t)} = \Delta [\Delta Z_{(t)}] = \Delta [Z_{(t)} - Z_{(t-1)}],$$

ou seja,

$$\Delta^2 Z_{(t)} = Z_{(t)} - 2Z_{(t-1)} + Z_{(t-2)}. \quad (3.2)$$

De modo geral, a n-ésima diferença de Z_t é

$$\Delta^n Z_{(t)} = \Delta [\Delta^{n-1} Z_{(t)}]. \quad (3.3)$$

Em situações normais, será suficiente tomar uma ou duas diferenças para que a série se torne estacionária.

3.2 Modelos para séries temporais

Segundo Morettin e Tolo (2006), os modelos utilizados para descrever séries temporais são processos estocásticos, isto é, processos controlados por leis probabilísticas e qualquer que seja a classificação que se faça para os modelos de série temporal, pode-se considerar um número muito grande de modelos diferentes para descrever o comportamento de uma série particular. Sobre isso, são abordadas as seguintes definições:

3.2.1 Processos estocásticos

Aqui tem-se uma definição e um exemplo gráfico dado por Morettin e Tolo (2006) sobre o que é um processo estocástico e como se comportam os dados observados.

Definição 3.1. (Processo estocástico): *Seja T um conjunto arbitrário. Um processo estocástico é uma família $Z = \{Z(t), t \in T\}$, tal que, para cada $t \in T$, $Z(t)$ é uma variável aleatória.*

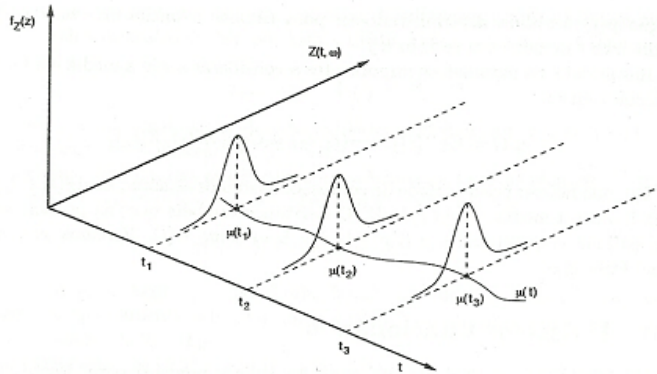


Figura 3: Um processo estocástico interpretado como uma família variáveis aleatórias.

Conforme a Figura 3, dada por Morettin e Toloi (2006), para cada $t \in T$, $Z(t)$ é uma variável aleatória definida sobre Ω , na realidade $Z(t)$ é uma função de dois argumentos, $Z(t, \omega)$, $t \in T$, $\omega \in \Omega$. A Figura 3 ilustra esta interpretação de processo estocástico, mostrando que para cada $t \in T$, $Z(t, \omega)$, temos uma variável aleatória $Z(t)$ com uma distribuição de probabilidade; é possível que a função densidade de probabilidade (fdp) no instante t_1 seja diferente da fdp no instante t_2 , para dois instantes t_1 e t_2 quaisquer, mas a situação usual é aquela em que a fdp de $Z(t, \omega)$ é a mesma para todo $t \in T$. Por outro lado, para cada $\omega \in \Omega$, fixado, obtém-se uma função de t , ou seja, uma realização ou trajetória do processo, ou ainda, uma série temporal.

Os processos estocásticos podem ser determinados através de classes que podem ser: processos estacionários ou não estacionários (de acordo com a independência ou não relativamente à origem dos tempos); processos normais (Gaussianos) ou não-normais (de acordo com as fdps que caracterizam os processos) e processos Markovianos ou não-Markovianos (de acordo com a independência dos valores do processo, em dado instante de seus valores em instantes precedentes), onde, estas classes na verdade são suposições simplificadoras para identificar modelos adequados em estudo com dados vindos de séries temporais.

3.2.2 Tipos de modelos

Os modelos para séries temporais são classificados segundo o número de parâmetros envolvidos, podendo ser:

1. Modelos paramétricos: caracterizam-se pelo número finito de parâmetros e a análise é feita no domínio de tempo. Os modelos mais utilizados são os modelos de erro (ou de regressão), os modelos autorregressivos e de médias móveis (ARMA), os modelos autorregressivos interados e de médias móveis (ARIMA), os modelos de memória longo (AFIRMA), modelos estruturais e modelos não-lineares;
2. Modelos não-paramétricos: caracterizam-se pelo infinito número de parâmetros. Os modelos mais utilizados são a função de autocovariância (ou autocorrelação) e sua transformada de Fourier, o espectro.

3.2.3 Função de autocorrelação (FAC)

Seja $\{X_t, t \in Z\}$ um processo real discreto, de média zero e função de autocovariância (FACV) $\gamma_\tau = E\{X_t X_{t+\tau}\}$. A FACV γ_τ satisfaz as seguintes propriedades:

1. $\gamma_0 > 0$,
2. $\gamma_{-\tau} = \gamma_\tau$,
3. $|\gamma_\tau| \leq \gamma_0$,
4. τ é não negativa definida, porque

$$\sum_{j=1}^n \sum_{k=1}^n a_j a_k \gamma_{\tau_j - \tau_k} \geq 0 \quad (3.4)$$

para quaisquer números reais a_1, \dots, a_n e τ_1, \dots, τ_n de Z .

A partir da FACV tem-se que, a função de autocorrelação (FAC) pode ser definida como:

$$\rho_\tau = \frac{\text{cov}(X_t, X_{t-\tau})}{\text{Var}(X_t)}, \quad (3.5)$$

onde, de acordo com Morettin e Tolo (2006) será escrita

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0}, \tau \in Z, \quad (3.6)$$

e tem as propriedades de γ_r exceto que agora $\rho_0 = 1$. Podendo ser estimada por

$$r_\tau = \hat{\rho}_\tau = \frac{\sum_{t=1}^{n-\tau} (X_t - \bar{X})(X_{t-\tau} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2}. \quad (3.7)$$

3.2.4 Modelos autorregressivos - AR(p)

São modelos cuja característica principal é que em lugar das variáveis independentes, o processo utilizar-se-á dos valores prévios da série temporal para estimação do modelo. Neste caso, a série temporal $Z_{(t)}$ é descrita somente pelos seus valores realizados e pelos termos aleatórios. No modelo Autorregressivo (AR), o comportamento de uma variável é explicado pelo seu próprio passado. Seja a_t um processo puramente aleatório com média μ e variância σ^2 , um processo Z_t é chamado de processo autoregressivo de ordem p , ou AR(p), se

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t, \quad (3.8)$$

Reescrevendo-se o modelo em termos de operador *backward*, temos o operador autorregressivo de ordem p :

$$\phi B = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, \quad (3.9)$$

onde,

$$\phi(B) \tilde{Z}_t = a_t. \quad (3.10)$$

3.2.5 Função de autocorrelação parcial (FACP)

A função de autocorrelação parcial é utilizada visando facilitar o processo de identificação do modelo. Segundo Morettin e Tolo (2006), denotando-se por ϕ_{kj} o j -ésimo coeficiente de um modelo AR(k), de tal modo que ϕ_{kk} seja o último coeficiente. Tem-se que

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

a partir das quais obtém-se as equações de Yule-Walker

$$\begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{bmatrix} \times \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix} \quad (3.11)$$

e em geral $\phi_{kk} = \frac{|P_k^*|}{|P_k|}$, em que P_k é a matriz de autocorrelações e P_k^* com a última coluna substituída pelo vetor de autocorrelações.

3.2.6 Modelos de médias móveis - MA(q)

Nestes modelos, a série temporal $Z(t)$ é resultado da combinação linear dos choques aleatórios (ruídos brancos) ocorridos no período corrente t e em períodos anteriores. Seja a_t um processo puramente aleatório com média μ e variância σ_a^2 . Um processo Z_t é chamado de média móvel de ordem q , ou MA(q), se

$$\tilde{Z}_t = \mu + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (3.12)$$

e sendo $\tilde{Z}_t = Z_t - \mu$, tem-se

$$\tilde{Z}_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t = \theta(B) a_t, \quad (3.13)$$

onde

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q. \quad (3.14)$$

é o operador de médias móveis de ordem q .

3.2.7 Modelos autorregressivos e de médias móveis - ARMA(p,q)

Para este modelo, a série temporal $Z(t)$ é função de seus valores históricos e pelos choques aleatórios correntes e passados. Os modelos autorregressivos são bastante populares em algumas áreas, como em Economia, onde é natural pensar o valor de alguma variável no instante t como função de valores defasados da mesma variável.

Segundo Morettin e Toloi (2006), para muitas séries encontradas na prática, se quiser-

mos um modelo com um número não muito grande de parâmetros, a inclusão de termos autoregressivos e de médias móveis é a solução adequada. A forma da combinação de modelos autorregressivos e de média móveis origina-se da seguinte forma

$$\hat{Z}_t = \phi_1 \hat{Z}_{t-1} + \dots + \phi_p \hat{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (3.15)$$

se $\phi(B)$ e $\theta(B)$ são os operadores autorregressivos e de médias móveis, respectivamente, pode ser escrito na forma compacta

$$\phi(B) \hat{Z}_t = \theta(B) a_t, \quad (3.16)$$

3.3 Modelos não estacionários

Segundo Ehlers (2007), os modelos AR, MA e ARMA, são apropriados para séries temporais estacionárias. Assim, para ajustar estes modelos a uma série temporal observada é necessário remover as fontes de variação não estacionárias. Por exemplo, se a série observada for não estacionária na média pode-se tentar remover a tendência tomando-se uma ou mais diferenças.

Sabe-se que, na prática, as séries comumente encontradas, especialmente as originadas do setor econômico, são não estacionárias, ou seja, comportam-se de formas diferentes em certos períodos de tempo. Para que seja possível utilizar um modelo Box e Jenkins será preciso recorrer a uma transformação, onde a série pode ser diferenciada para tornar-se estacionária, como mostra o exemplo dado por Morettin e Toloi (2006), onde tem-se um Z_t não estacionário mas

$$W_t = Z_t - Z_{t-1} = (1 - B) Z_t = \Delta Z_t$$

é estacionário.

3.3.1 Modelo autorregressivo integrado de médias móveis - ARIMA (p,d,q)

O modelo ARIMA (p,d,q) é um caso especial de um processo integrado. Em geral, diz-se que Z_t é integrado de ordem d se $\Delta^d Z_t$ for estacionário, e escreve-se $Z_t \sim I(d)$. Se $d=0$, Z_t é estacionário (MORETTIN; TOLOI, 2006).

Se $W_t = \Delta^d Z_t$ for estacionário, pode-se representar W_t por um modelo ARMA (p,q), ou seja,

$$\phi(B) W_t = \theta(B) a_t. \quad (3.17)$$

Se W_t for uma diferença de Z_t , então Z_t é uma integral de W_t , daí diz-se que Z_t segue um modelo autorregressivo, integrado, de médias móveis, ou modelo ARIMA,

$$\phi(B) \Delta^d Z_t = \theta(B) a_t, . \quad (3.18)$$

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

Na prática valores pequenos são em geral especificados para d , sendo $d = 1$ o valor mais frequentemente utilizado e excepcionalmente $d = 2$. Sobre isso Morettin e Tolo (2006) afirmam que nesses dois casos de não estacionariedade homogênea, a série pode se comportar das seguintes formas:

- (a) séries não estacionárias quanto ao nível: oscilam ao redor de um nível médio durante algum tempo e depois saltam para outro nível temporário. Para torná-la estacionária é suficiente tomar uma diferença; este é o caso típico de séries econômicas, como mostra a Figura 4 dada por Fischer (1982).

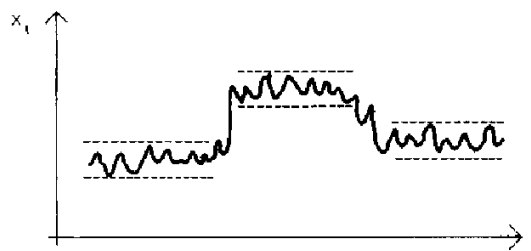


Figura 4: Série não estacionária quanto ao nível.

- (b) séries não estacionárias quanto a inclinação: oscilam numa direção por algum tempo e depois mudam para outra direção temporária. Para torná-la estacionária é necessário tomar a segunda diferença (como observado na Figura 2).

Vale notar, que para dados reais, um modelo ARIMA (e de fato qualquer modelo) é no máximo uma aproximação para o verdadeiro processo gerador dos dados. Na prática

pode ser bem difícil distinguir entre um processo estacionário com memória longa e um processo não estacionário homogêneo (EHLERS, 2007).

3.3.2 Identificação do modelo

A identificação compreende um conjunto de procedimentos com o objetivo de proporcionar uma ideia aproximada da estrutura do modelo. Assim, a identificação é a etapa onde se estabelecem as características básicas da série, tais como o tipo de processo gerador (estacionário ou não estacionário); a existência ou não das componentes de sazonalidade e tendência, e que capacita o pesquisador a estabelecer a ordem do modelo que se ajusta a série observada; e as estimativas iniciais para os valores dos parâmetros envolvidos (FISCHER, 1982).

Segundo Morettin e Tolo (2006), o objetivo da identificação é determinar os valores de p , d e q do modelo ARIMA (p,d,q) , além de estimativas preliminares dos parâmetros a serem usados no estágio de estimação. Esta etapa é de fundamental importância e também é considerada uma das mais difíceis, podendo, para este fim, utilizar-se dos seguintes facilitadores:

- (i) A FAC e a FACP: realiza-se a identificação dos modelos AR, MA e ARMA através das características da FAC e FACP, utiliza-se a Tabela 1.

Tabela 1: Identificação dos modelos ARIMA.

Tipo de processo	FAC	FAC Parcial
AR(p)	declinante	truncada em p
MA(q)	truncada em q	declinante
ARMA(p,q)	declinante	declinante

Um processo será AR(p) quando apresentar FAC declinante e FAC Parcial truncada em p , será MA(q) ao apresentar FAC truncada em q e FAC Parcial declinante, ou ainda será um ARMA(p, q) quando apresentar FAC e FAC Parcial declinante (MORAIS, 2010).

Além da verificação do valores de p e q , será necessário identificar o grau das diferenças d para modelos do tipo ARIMA(p, d, q).

- (ii) Critérios de Informação: Em muitos casos, no estudo de séries temporais, é possível encontrar vários modelos adequados no que se refere ao comportamento dos resíduos.

Uma forma de “discriminar” entre estes modelos competidores é utilizar os chamados critérios de informação que levam em conta não apenas a qualidade do ajuste, mas também penalizam a inclusão de parâmetros extras. Assim, um modelo com mais parâmetros pode ter um ajuste melhor mas não necessariamente será preferível em termos de critério de informação (EHLERS, 2007). Pretende-se selecionar o modelo cujo critério de informação calculado seja mínimo.

A regra mais utilizada em séries temporais é o chamado critério de informação de Akaike, denotado por AIC e geralmente é dado por

$$AIC = -2 \log L + 2m, \quad (3.19)$$

em que m é o número de parâmetros e L a função de máxima verossimilhança. Para dados distribuídos normalmente e utilizando-se estimativas de máxima verossimilhança dos parâmetros, tem-se que

$$AIC = n \log (\hat{\sigma}_\varepsilon^2) + 2m, \quad (3.20)$$

onde $\hat{\sigma}_\varepsilon^2 = (1/n) \sum \hat{\varepsilon}_t^2$.

A presença de p e q , na fórmula da estatística AIC, tem por objetivo “penalizar” os modelos com muitos parâmetros, tendo em vista que modelos mais parcimoniosos devem ser privilegiados, por apresentarem menor número de parâmetros a serem estimados. A comparação de modelos também poderá ser realizada com o Critério de Informação Bayesiano (BIC) como aproximações do fator de Bayes. O BIC é basicamente uma modificação do AIC, na forma de penalizar a inclusão de parâmetros extras e é dado por

$$BIC = -2 \log L + m \log n. \quad (3.21)$$

Igualmente ao critério AIC, o modelo mais adequado será aquele que tiver o menor valor da estatística BIC.

Observa-se que este critério de informação penaliza bem mais a inclusão de parâmetros do que o AIC e portanto tende a selecionar modelos mais parcimoniosos (EHLERS, 2007).

3.3.3 Estimação do modelo

Conforme aborda Bezerra (2006), as estimativas do modelo, dadas pelo método de máxima verossimilhança é vantajoso, pois sobre certas condições gerais, muitos resultados já são conhecidos, para o caso de grandes amostras, além do que, todas as informações contidas nos dados são utilizadas, entretanto, uma desvantagem é que para os primeiros valores de t , deve-se trabalhar especificamente com a função densidade de probabilidade (f.d.p) conjunta.

Para um dado conjunto de observações Z_1, \dots, Z_n a função de verossimilhança L é definida como a f.d.p. conjunta dos dados observados, em função dos parâmetros do modelo. Para os modelos ARIMA, L será uma função de $\phi's, \theta's, \mu$ e σ_a^2 dadas as observações Z_1, \dots, Z_n . Os estimadores de máxima verossimilhança são definidos como aqueles valores dos parâmetros, com relação aos dados observados, que são os mais verossímeis, e maximizam a função de verossimilhança.

3.3.4 Diagnóstico do modelo

Após a estimação do modelo, se faz necessário averiguar a adequação do mesmo, ou seja, se o modelo será bem representativo aos dados da série estudada. Ao invés de olhar para as autocorrelações residuais individualmente, pode-se testar se um grupo de autocorrelações é significativamente diferente de zero (EHLERS, 2007). Entre os métodos utilizados para esta avaliação, está a estatística Q , calculada pelos testes de Box-Pierce ou Ljung-Box que avaliam a qualidade do modelo, em que verificam se o conjunto de autocorrelações residuais é estatisticamente diferente de zero.

$$Q = n \sum_{k=1}^m \rho_k^2. \quad (3.22)$$

Na prática o número m de autocorrelações amostrais é tipicamente escolhido entre 15 e 30 (EHLERS, 2007). Se o modelo for apropriado, Q terá uma distribuição aproximadamente Qui-quadrado com $m - p - q$ graus de liberdade.

Contudo, o teste Box-Pierce não tem bom desempenho em amostras pequenas pois a distribuição afasta-se da Qui-quadrado, então aplica-se o teste Ljung-Box, que é aproximadamente Qui-quadrado com $m - p - q$ graus de liberdade.

$$Q = n(n+2) \sum_{k=1}^m \frac{\hat{\rho}_k^2}{n-k}. \quad (3.23)$$

Em síntese, pode-se explicar que a estatística Q , calculada por meio do teste Box-Pierce ou pelo teste de Ljung-Box é utilizada para verificar conjunto de autocorrelações de resíduos sendo estatisticamente diferente ou não de zero. Assim, ambos os testes possuem distribuição χ^2 (Qui-quadrado), de modo que, calcula-se o valor de Q , e conclui-se da seguinte forma: se $Q_{\text{calculado}} > Q_{\text{tabelado}}$, então rejeita-se a hipótese nula (H_0) de que os resíduos não são correlacionados, ou seja, pelo menos uma autocorrelação difere estatisticamente de zero, com isso, o modelo deverá ser rejeitado.

É importante ressaltar que, ao estimar-se um modelo, deseja-se que o erro produzido por ele possua característica de ruído branco (independente e identicamente distribuído). Segundo Morettin e Tolo (2006), a hipótese de ruído branco para os resíduos é rejeitada para valores grandes $Q(K)$. Em geral basta utilizar as 10 ou 15 primeiras $\hat{\rho}_k$.

3.3.5 Previsão do modelo

Segundo Morettin e Tolo (2006), designando-se $\hat{Z}_t(h)$, a previsão de origem t e horizonte h , $\hat{Z}_t(h)$ poderá ser expressado de três maneiras:

(a) Previsão utilizando a equação de diferenças

Tomando a esperança condicional em uma equação de diferenças, tem-se

$$\hat{Z}_t(h) = \varphi_1[Z_{t+h-1}] + \dots + \varphi_{p+d}[Z_{t+h-p-d}] - \theta_1[a_{t+h-1}] - \dots - \theta_q[a_{t+h-q}] + [a_{t+h}] \quad (3.24)$$

para $h \geq 1$. Aqui deve-se utilizar os seguintes fatos

$$\begin{aligned} [Z_{t+k}] &= \hat{Z}_t(h), & k > 0 \\ [Z_{t+k}] &= Z_{t+k}, & k \leq 0 \\ [a_{t+k}] &= 0, & k > 0 \\ [a_{t+k}] &= a_{t+k}, & k \leq 0 \end{aligned}$$

(b) Previsão utilizando a forma de choques aleatórios

Utilizando-se a equação de choques aleatórios, tem-se

$$\hat{Z}_t(h) = \psi_1 [a_{t+h-1}] + \dots + \psi_{h-1} [a_{t+1}] + \psi_h [a_t] + \dots + [a_{t+h}] \quad (3.25)$$

(c) Previsão utilizando a forma invertida

Obtida a partir da equação da forma invertida

$$\hat{Z}_t(h) = \sum_{j=1}^{\infty} \pi_j [Z_{t+h-j}] + [a_{t+h}]. \quad (3.26)$$

A equação de previsão, considerada como uma função de h , com origem t fixa, satisfaz a equação de diferenças

$$\sum_{i=1}^{p+d} \varphi_i \hat{Z}_t(h-1), h > q$$

ou

$$\varphi(B) \hat{Z}_t(h) = (1-B)^d \phi(B) \hat{Z}_t(h) = 0, h > q$$

com $\varphi(B)$ operando sobre h .

3.3.6 Medidas de precisão para previsão do modelo

De acordo com Hyndman e Koehler (2005), denotando-se Y_t como a observação no tempo t e F_t a previsão de t . Em seguida, definindo-se o erro de previsão $e_t = Y_t - F_t$. As previsões podem ser computadas a partir de uma base de tempo comum, e ser de diferentes horizontes de previsão. Assim, pode-se calcular as previsões F_{n+1}, \dots, F_{n+m} baseado em dados de tempo $t = 1, \dots, n$. Alternativamente, as previsões podem ser de diferentes tempos de base, e ser de um horizonte de previsão consistente. Ou seja, pode-se calcular as previsões F_{1+h}, \dots, F_{m+h} , onde cada F_{j+h} é baseado em dados de tempo $t = 1, \dots, j$. As previsões de amostra nos exemplos acima basearam-se no segundo cenário com $h = 1$. Um terceiro cenário surge quando deseja-se comparar a precisão dos métodos através de muitas séries em um único horizonte de previsão. Em seguida, calculamos um único F_{n+h} com base em dados de tempos $t = 1, \dots, n$ para cada uma das m séries diferentes.

Existem algumas medidas de precisão geralmente usadas, cuja escala depende da escala dos dados. Estes são úteis na comparação de métodos diferentes com o mesmo conjunto de dados, mas não deve ser usado, por exemplo, entre os conjuntos de dados que

têm escalas diferentes. As escalas de medidas dependentes mais utilizadas baseiam-se no erro absoluto ou erros quadrados:

(i) Erro quadrático médio (MSE) = média (e_t^2)

(ii) Raiz quadrada do erro quadrático médio (RMSE) = \sqrt{MSE}

(iii) Erro médio absoluto (MAE) = média ($|e_t|$)

(iv) Erro mediano absoluto (MdAE) = mediana ($|e_t|$)

Muitas vezes, o RMSE é preferido ao MSE que seja na mesma escala como os dados. Historicamente, o RMSE e MSE foram populares, em grande parte devido a sua relevância teórica em modelagem estatística. No entanto, eles são mais sensíveis a *outliers* que MAE ou MdAE.

3.4 Modelos Sazonais

Segundo Ehlers (2007), muitas séries temporais contém uma componente periódica sazonal que se repete a cada s observações ($s > 1$). Por exemplo, com dados mensais e $s = 12$ tipicamente espera-se que Z_t dependa de Z_{t-12} e talvez de Z_{t-24} além de Z_{t-1}, Z_{t-2}, \dots . Neste caso, tomar a primeira diferença $z_t - z_{t-1}$ não é suficiente para tornar a série (aproximadamente) estacionária. A forma apropriada de diferenciar dados com padrão sazonal acentuado é tomar diferenças no período sazonal. Por exemplo, para dados mensais a primeira diferença sazonal é

$$\nabla_{12}z_t = (1 - B^{12})z_t = z_t - z_{t-12}; \quad (3.27)$$

e terá variabilidade menor do que a primeira diferença não sazonal $\nabla z_t = z_t - z_{t-1}$, sendo portanto mais fácil de identificar e estimar.

Em geral, uma diferença sazonal é denotada por ∇_s onde s é o período sazonal. A D -ésima diferença sazonal é então denotada por ∇_s^D . Combinando-se os dois tipos de diferenciação, obtém-se o operador $\nabla^d \nabla_s^D$. Para se obter “d” e “D”, deve-se utilizar, conjuntamente, os operadores $\nabla^d = (1 - B)^d$ e $\nabla_s^D = (1 - B^s)^D$ de maneira a tornar $W_t = \nabla^d \nabla_s^D Z_t$ estacionária (FISCHER, 1982).

Por exemplo, tomando-se 1 diferença simples e 1 sazonal em uma série mensal tem-se que

$$\nabla\nabla_{12}z_t = z_t - z_{t-1} - z_{t-12} + z_{t-13}; \quad (3.28)$$

3.4.1 Modelos SARIMA

Para Ehlers (2007), a generalização do modelo ARIMA para lidar com sazonalidade é definida por meio de um modelo ARIMA sazonal multiplicativo, denominado SARIMA, dado por

$$\phi(B)\Phi(B^s)W_t = \theta(B)\Theta(B^s)\varepsilon_t; \quad (3.29)$$

em que

$$\begin{aligned} \phi(B) &= (1 - \alpha_1 B - \dots - \alpha_p B^p) \\ \Phi(B^s) &= (1 - \phi_s B^s - \dots - \phi_p B^{Ps}) \\ W_t &= \nabla^d \nabla_s^D Z_t \\ \theta(B) &= (1 + \beta_1 B + \dots + \beta_q B^q) \\ \Theta(B^s) &= (1 + \theta_s B^s + \dots + \theta_Q B^{Qs}) \end{aligned}$$

conforme Ehlers (2007), este modelo é chamado SARIMA multiplicativo de ordem $(p, d, q) \times (P, D, Q)_s$ e parece extremamente complicado à primeira vista, mas na prática, os valores de d e D em geral não serão maiores do que 1 e um número pequeno de coeficientes será suficiente. Por exemplo, com $P = 1$ tem-se que

$$\Phi(B^s) = (1 - \alpha_s B^s)$$

que significa simplesmente que W_t depende de W_{t-s} . A série W_t é formada a partir da série original tomando-se diferenças simples para remover a tendência e diferenças sazonais para remover a sazonalidade. O modelo SARIMA $(1, 0, 0) \times (0, 1, 1)_{12}$ para dados mensais, ou seja, tem-se um termo autorregressivo e um termo média móvel sazonal modelando a primeira diferença sazonal. Conforme Fischer (1982), SARIMA é a simbologia para um modelo ARIMA sazonal multiplicativo de ordem $(p, d, q) \times (P, D, Q)$. O modelo pode ser escrito como

$$(1 - \alpha B)(1 - B^{12})Z_t = (1 - \theta B^{12})\varepsilon_t$$

e desenvolvendo os produtos obtém-se que

$$Z_t = Z_{t-12} + \alpha (Z_{t-1}, Z_{t-13}) + \varepsilon_t + \theta \varepsilon_{t-12}.$$

Assim, Z_t depende de Z_{t-1} , Z_{t-12} e Z_{t-13} além do erro no tempo $t-12$. Para finalizar, ao ajustar um modelo sazonal aos dados, a primeira tarefa é especificar os valores de d e D que tornam a série (aproximadamente) estacionária e remove a maior parte da sazonalidade. Posteriormente os valores de p , P , q e Q devem ser especificados com base nas funções de autocorrelação e autocorrelação parcial da série diferenciada. Os valores de P e Q são especificados basicamente a partir de rk , $k = s, 2s, \dots$. Por exemplo, para dados mensais, se $r12$ é grande, mas $r24$ é pequeno, isto sugere que um termo média móvel sazonal pode ser adequado. Após ter identificado, por tentativa, o que parece ser um modelo SARIMA razoável, os parâmetros serão estimados por algum procedimento iterativo similar àqueles propostos para modelos ARMA.

3.5 Metodologia de Box e Jenkins

Segundo Fischer (1982), o tratamento desenvolvido por Box & Jenkins para a análise das séries estocásticas de tempo está baseado no fato de que, embora seus respectivos valores no tempo x , apresentem correlação serial, cada um deles pode ser considerado como gerado por uma sequência de choques “ a_t ”, $t \in T$, aleatórios e independentes entre si, cada um possuindo uma determinada distribuição, com $\mu_{(t)} = 0$ e variância constante σ_a^2 .

A sequência de choques aleatórios a_t , $t \in T$, com as características acima mencionadas é denominada de processo de ruído branco.

O fundamento da metodologia de Box & Jenkins, para a busca de uma classe geral de modelos capazes de representar o processo gerador da série estocástica, está assentado no teorema da decomposição de Wold para séries estacionárias. Esse teorema demonstra que todo o processo estocástico estacionário pode ser decomposto em um modelo linear de tipo média móvel, cuja formulação é

$$Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \dots, \quad (3.30)$$

onde μ e ψ_j , $j=1,2,3,\dots$ são parâmetros estabelecidos pelos valores da série, Z_t é a observação no período “ t ”, a_t o ruído branco no período “ t ” e os ψ_j devem formar uma

série absolutamente convergente para que Z seja estável, isto é, tenha variância finita e capaz de ser somável e mantenha a condição de estacionariedade.

O valor assumido por μ (onde, μ é a constante sazonal “médias mensais”) determina o ajuste aplicado aos dados. Quanto menor o valor da constante, mais estável serão as previsões, visto que a utilização de baixo valor implica na atribuição de peso maior as observações passadas e, conseqüentemente, qualquer flutuação aleatória no presente contribui com menor importância para a obtenção da previsão. Contudo não há metodologia que oriente quanto a seleção de um valor apropriado para μ , sendo normalmente encontrado por tentativa e erro. Um procedimento mais objetivo seria a seleção do valor que forneça a melhor previsão das observações contidas na série temporal (CARVALHO, 2007).

Conforme Morettin e Tolo (2006), essa metodologia consiste em ajustar modelos autorregressivos integrados de médias móveis, ARIMA (p,d,q), a um conjunto de dados, em que :

p= número de parâmetros autoregressivos;

d= número de diferenças aplicadas na série;

q= número parâmetros de médias móveis.

Em que, os modelos ARIMA são resultados da combinação de três componentes denominados “filtros”: o componente autorregressivo (AR), o filtro de integração (I) e o componente de médias móveis (MA). A estratégia para a construção do modelo será baseada em um ciclo iterativo, no qual a escolha da estrutura do modelo é baseada nos próprios dados. Os estágios do ciclo iterativo são:

- (i) *Especificação*: uma classe geral de modelo é considerada para a análise;
- (ii) *Identificação*: feita com base na análise de autocorrelações, autocorrelações parciais e outros critérios;
- (iii) *Estimação*: na qual os parâmetros do modelo identificado são estimados;
- (iv) *Verificação*: realizada através de uma análise de resíduo, para se saber se este é adequado para os fins em vista (previsão por exemplo).
- (v) *Validação*: feito com base nos pressupostos de adequação do modelo.

Caso o modelo não seja adequado o ciclo é repetido, voltando-se à fase de identificação. Um procedimento muitas vezes utilizado, é identificar vários modelos, que são estimados e verificados.

O conhecimento prévio de algumas notações de operadores são importantes para um melhor manejo dos modelos a serem estudados. Segundo Morettin e Tolo (2006), estes operadores são:

a) operador translação para o passado, denotado por B e definido por

$$BZ_t = Z_{t-1}, B^m Z_t = Z_{t-m}; \quad (3.31)$$

b) operador translação para o futuro, denotado por F e definido por

$$FZ_t = Z_{t+1}, F^m Z_t = Z_{t+m}; \quad (3.32)$$

c) operador diferença, já definido antes

$$\Delta Z_t = Z_t - Z_{t-1} = (1 - B) Z_t; \quad (3.33)$$

segue-se que

$$\Delta = 1 - B,$$

d) operador soma, denotado por S e definido por

$$SZ_t = \sum_{j=0}^{\infty} Z_{t-j} = Z_t + Z_{t-1} + \dots = (1 + B + B^2 + \dots) Z_t, \quad (3.34)$$

do que segue

$$SZ_t = (1 - B)^{-1} Z_t = \Delta^{-1} Z_t,$$

ou seja,

$$S = \Delta^{-1}.$$

4 Aplicação

Os dados que compõem a série em estudo foram obtidos através do site Banco Central do Brasil (2012) e referem-se a quantidade de cheques que foram compensados no Estado da Paraíba no período de janeiro 1997 a fevereiro de 2012, em um total de 182 observações, contudo, no estudo desta série, serão utilizadas 168 observações para o ajuste do modelo, perfazendo o período de Janeiro 1997 a Dezembro de 2010. As 14 observações restantes, foram utilizadas para validação cruzada do modelo com o auxílio do software R versão 2.14.2 (R development core team, 2011).

A Figura 5, mostra o comportamento da série de cheques compensados seguindo uma não estacionariedade decrescente ao longo do tempo, e apresenta-se com certa sazonalidade constante no intervalo de cada 12 meses, evidenciando ainda, através da representação de tendência que a compensação de cheques decai rapidamente após o ano 1998, retoma um crescimento a partir do ano 2000 até o ano de 2002, de onde tem início um decréscimo mais lento. Esse comportamento em declínio na movimentação de cheques, está relacionado a utilização dos cartões de crédito na aquisição de bens e serviços. Como aborda Carvalho (2007) em estudo similar, o mercado brasileiro de cartões, neste caso incluindo os cartões de crédito, débito e os de loja, vem apresentando um crescimento superior aos índices do PIB. O ano de 2006 encerrou com mais de 390 milhões de cartões, um aumento de 15% em comparação a 2005, movimentando aproximadamente R\$ 261,8 milhões, um crescimento de 24% segundo dados da Associação Brasileira das Empresas de Cartões de Crédito e Serviços (ABECS), mas que distribuídos entre toda a população brasileira chega à média de apenas dois cartões por habitante no Brasil (A tendência é um aumento ainda consistente dos cartões, considerando o potencial do crescimento do Brasil). No aspecto macroeconômico os cartões contribuem para a modernização do sistema de pagamentos no varejo, uma vez que apresentam maior eficiência com relação aos instrumentos em papel, como é o caso dos cheques. As empresas de cartões reconhecem que por questões de segurança, facilidades na organização do orçamento doméstico e benefícios como milhagens, a cada dia o consumidor prefere mais o cartão, em detrimento ao cheque.

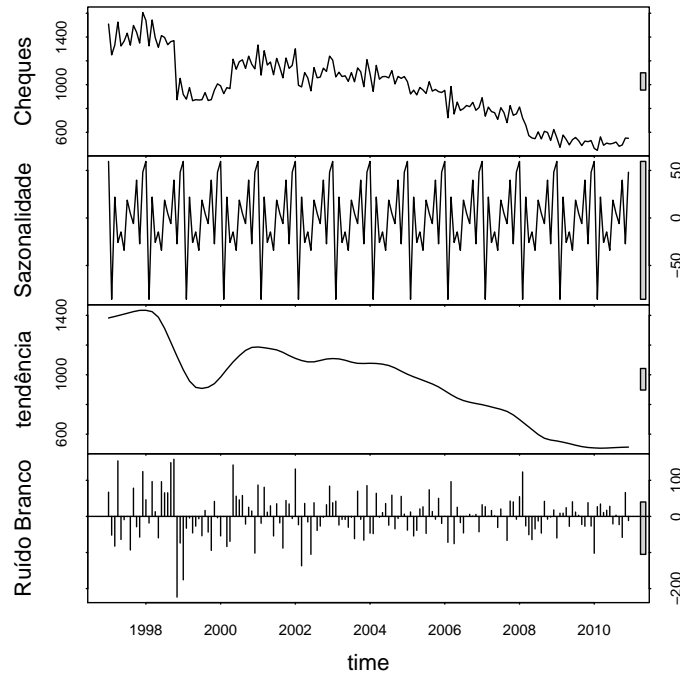
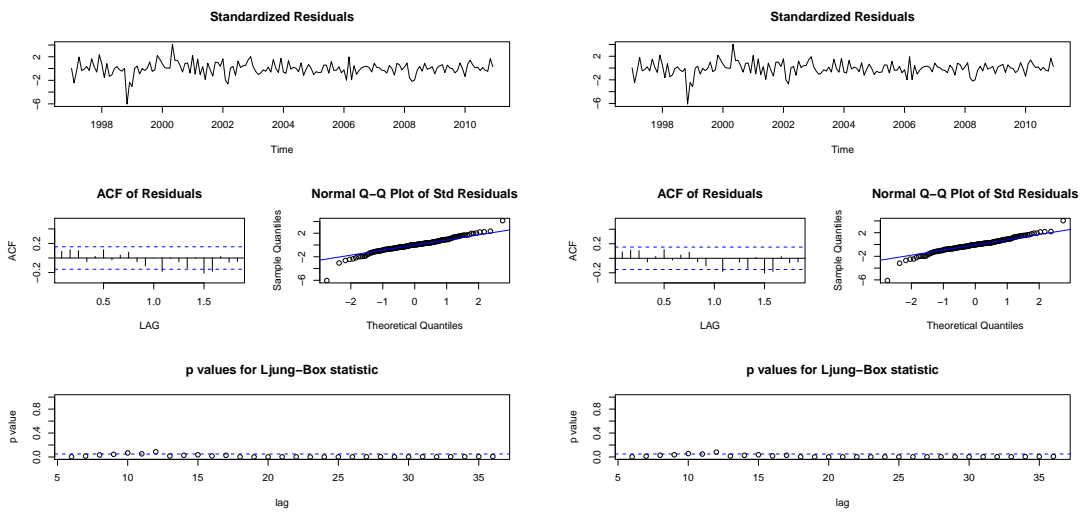


Figura 5: Cheques compensados na Paraíba no período de Janeiro 1997 a Dezembro de 2010.

A função *auto.arima*, disponível no pacote *forecast* implementado no *software* R, seleciona um modelo com base na estatística BIC. O modelo gerado foi um modelo sarima de ordem $(2,1,0) \times (2,0,1)$. Supõe-se, que este seja o modelo que melhor explique os dados de cheques compensados, no entanto, sob a hipótese nula, de que, os erros são ruído branco, o teste Ljung-Box, ilustrado nas Figuras 6(a) e 6(b) do referido modelo, com e sem constante, rejeita-se H_0 , indicando que os erros não são ruído branco, pois o valor $P < 0,05$, logo os modelos não seriam bem representativos para o conjunto de dados.

O modelo sugerido pelo processo *auto.arima* foi ajustado por meio de aproximação, pois o máximo da função de verossimilhança não foi encontrado de forma exata, deste modo, foi ajustado um novo modelo por meio da função *auto.arima*, com uma restrição na função (*approximation = FALSE*), que força com que o programa ajuste o modelo e encontre o estimador de máxima verossimilhança de forma exata, o modelo ajustado foi um sarima $(0,1,2) \times (0,0,1)$, que foi logo descartado devido aos resíduos não serem ruído branco (Valor $P < 0,05$ pelo teste de Ljung Box) e o $BIC = 1976,08$ e o $AIC = 1963,61$ serem muito altos comparados a outros modelos ajustados.

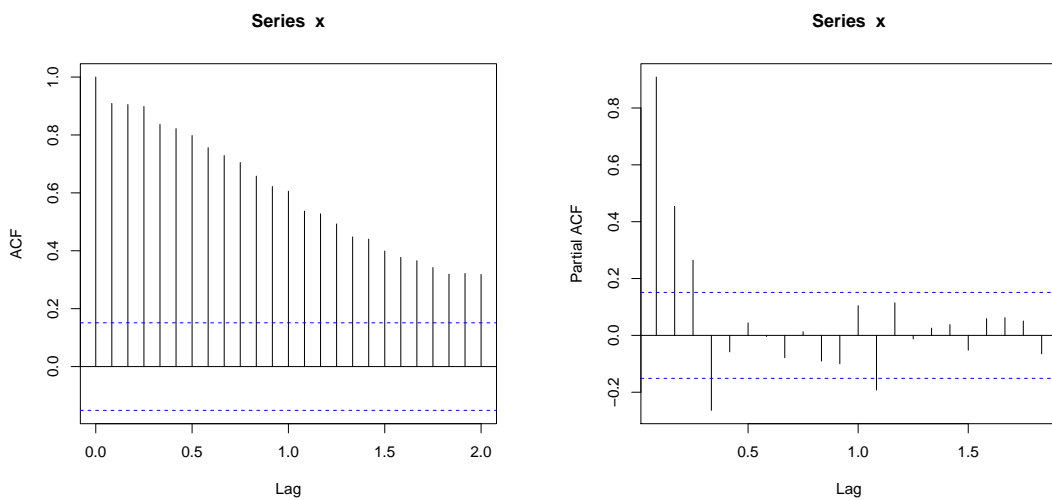
Tendo em vista estas dificuldades, partiu-se para os métodos usuais para ajuste de modelos de forma empírica. As funções de autocorrelação e autocorrelação parcial serão



(a) Modelo sugerido pelo auto.arima com constante (b) Modelo sugerido pelo auto.arima sem constante

Figura 6: Teste Ljung-Box para validação do modelo.

de fundamental importância para eleger possíveis modelos apropriados à série estudada. Observando os gráficos 7(a) e 7(b) é possível indentificar por meio de defasagem, um comportamento em declínio em ambos e sugere-se o modelo ARIMA. Levando-se em consideração o comportamento sazonal, serão propostos modelos sarima multiplicativos, buscando compará-los, afim de eleger o melhor possível.



(a) Gráfico de autocorrelação (b) Gráfico autocorrelação parcial

Figura 7: Gráficos para selecionar possíveis modelos adequados.

Inicialmente, optou-se, pela escolha, em função dos critérios, que geralmente são muito utilizados, como, o BIC e o AIC. Sabe-se que o uso do critério Bayesiano busca o menor valor possível, possibilitando modelos mais parcimoniosos, que neste caso, como pode-se

ver na Tabela 2, foi um BIC=1955.65 e um AIC=1936.94, o que nos leva a indicar que o modelo ajustado mais adequado é o M12. Entretanto, como o maior objetivo deste estudo é a previsão, tomou-se então como parâmetro fundamental de interesse, a busca pelo modelo que satisfaça critérios de previsão.

Partiu-se então, para eleição do modelo com base na medida de precisão MAE, já que o RMSE é mais sensível a *outliers*, com isso, o modelo Sarima M4 de ordem $(3,1,0) \times (2,0,2)$, com uso da constante, é designado como o modelo mais adequado para representar o conjunto de dados em estudo, uma vez que teoricamente, dará uma boa exatidão aos valores preditos. Vale salientar, que o uso da constante no modelo, determina o ajuste aplicado aos dados. Quanto menor o valor da constante, mais estável serão as previsões. O ajuste do modelo com melhor MAE, nos leva a um modelo com um BIC moderado, a diferença do BIC do modelo selecionado M4 e o modelo utilizado em comparação com os outros modelos foi menor que 10 unidades.

Tabela 2: Modelos associados aos respectivos critérios de seleção.

Modelo	AIC	BIC	$\lambda(\theta)$	MAE	RMSE
M1 (0 1 2) (0 0 1)*	1964.6	1980.19	977.3	58.7161	83.7061
M2 (2 1 0) (2 0 1)*	1945.9	1967.73	965.95	52.2869	75.2854
M3 (3 1 1) (2 0 1)*	1942.13	1970.19	962.06	51.6579	73.8221
M4 (3 1 0) (2 0 2)*	1939	1967.06	960.5	48.8770	69.8661
M5 (3 1 0) (1 0 1)*	1938.53	1960.36	962.27	51.2555	73.4885
M6 (3 1 1) (1 0 1)*	1940.3	1965.24	962.15	51.3726	73.6867
M7 (2 1 2) (2 0 1)*	1940	1968.06	961	52.2891	73.6588
M8 (0 1 2) (0 0 1)**	1963.61	1976.08	977.81	58.7161	83.9511
M9 (2 1 0) (2 0 1)**	1944.62	1963.33	966.31	52.2248	75.1864
M10 (3 1 1) (2 0 1)**	1940.51	1965.46	962.26	51.7892	73.6831
M11 (3 1 0) (2 0 2)**	1939.24	1964.19	961.62	49.2227	70.9997
M12 (3 1 0) (1 0 1)**	1936.94	1955.65	962.47	51.3399	73.4154
M13 (3 1 1) (1 0 1)**	1938.67	1960.5	962.34	51.5616	73.5933
M14 (2 1 2) (2 0 1)**	1938.41	1963.35	961.21	52.3267	73.6431

Nota: * com constante; ** sem constante; $\lambda(\theta)$ log verossimilhança.

Após identificar o modelo pelo critério MAE, observa-se na Figura 8, que o teste Ljung-Box do modelo eleito, indica, que o grupo de autocorrelações estão acima do limiar de 5%, logo a escolha foi apropriada, pois a estatística de Ljung-Box, é um teste estatístico de hipótese nula de ausência de autocorrelação. Assim, se para toda defasagem a estatística, podemos considerar que os resíduos se comportam como ruído branco e que, portanto, o modelo ajustado é adequado. O modelo escolhido apresenta as seguintes estimativas dos parâmetros que o compõe conforme Tabela 3.

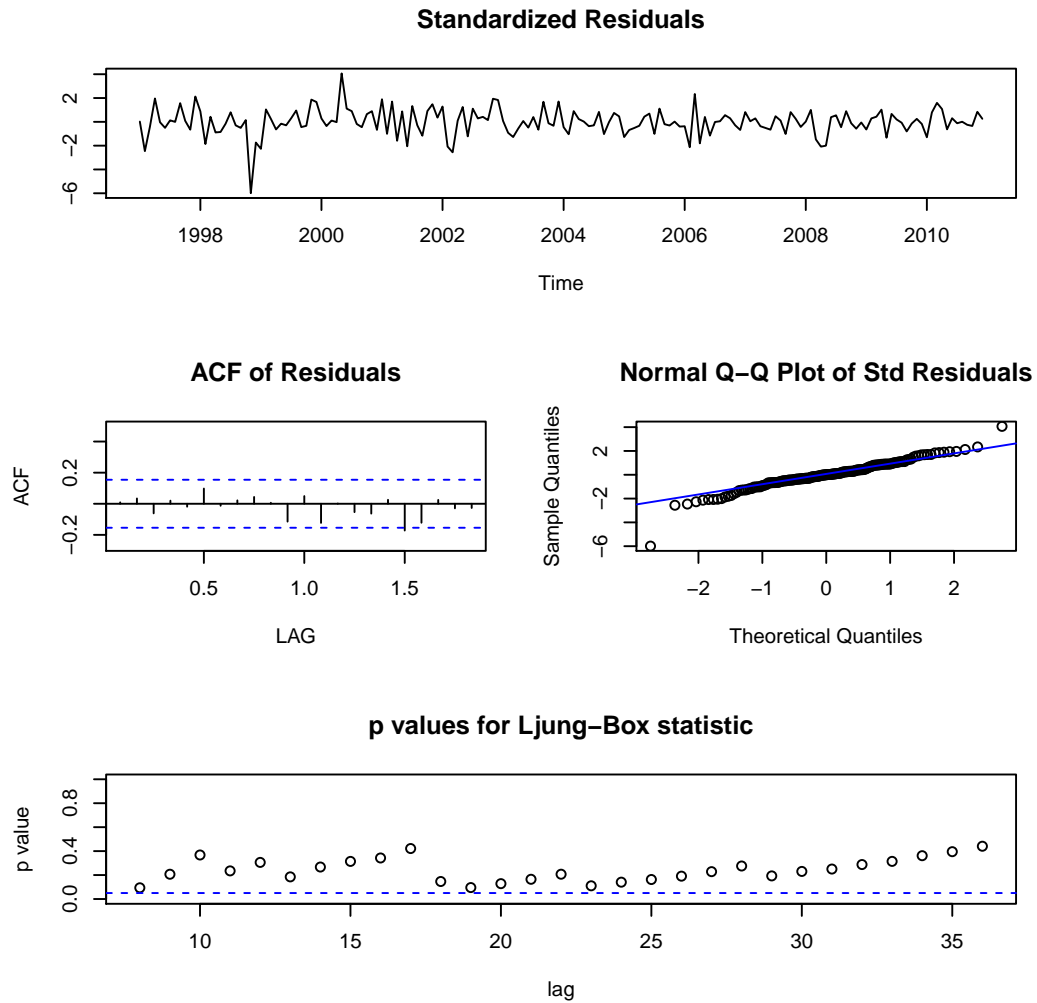


Figura 8: Gráfico para validação do modelo.

Tabela 3: Estimativas e Erros dos parâmetros do modelo selecionado.

Estadísticas	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\Phi}_1$	$\hat{\Phi}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\mu}$
Estimativas	-0.6040	-0.2352	0.2104	1.7741	-0.8132	-1.8977	0.9997	-5.4294
Erro Padrão	0.0767	0.0880	0.0777	0.1192	0.1057	0.2685	0.2754	7.2469

Observa-se na Tabela 4, os valores preditos para os quatorze meses consecutivos, ao término do período da série de cheques compensados em estudo, visando previsões com 80% e 95% de confiança.

Para uma melhor visualização, a Figura 9 apresenta as previsões para os 14 meses, utilizados para validação (em laranja 95% e amarelo 80%).

Na tabela 5, tem-se o erro de previsão, para cada uma das 14 observações, por meio do erro de previsão pode-se calcular o teste Shapiro-Wilk, com intuito de checar se os

Tabela 4: Previsão para quatorze meses com intervalo de confiança.

Mês/ano	predição	Li 80	Ls 80	Li 95	Ls 95
jan 2011	543.18	448.85	637.51	398.91	687.44
fev 2011	466.37	364.94	567.8	311.25	621.49
mar 2011	504.54	391.71	617.37	331.98	677.09
abr 2011	434.43	298.85	570	227.09	641.77
mai 2011	478.48	335.67	621.28	260.08	696.88
jun 2011	425.59	270.83	580.35	188.91	662.27
jul 2011	456.33	289.14	623.53	200.63	712.03
ago 2011	467.18	292.31	642.05	199.74	734.62
set 2011	428.79	243.69	613.9	145.7	711.89
out 2011	466.44	272.27	660.62	169.48	763.41
nov 2011	462.6	260.76	664.43	153.91	771.28
dez 2011	469.3	258.84	679.76	147.43	791.16
jan 2012	467.13	249.77	684.48	134.7	799.55
fev 2012	385.63	161.36	609.9	42.64	728.63

Nota: Li (limite inferior); Ls (limite superior).

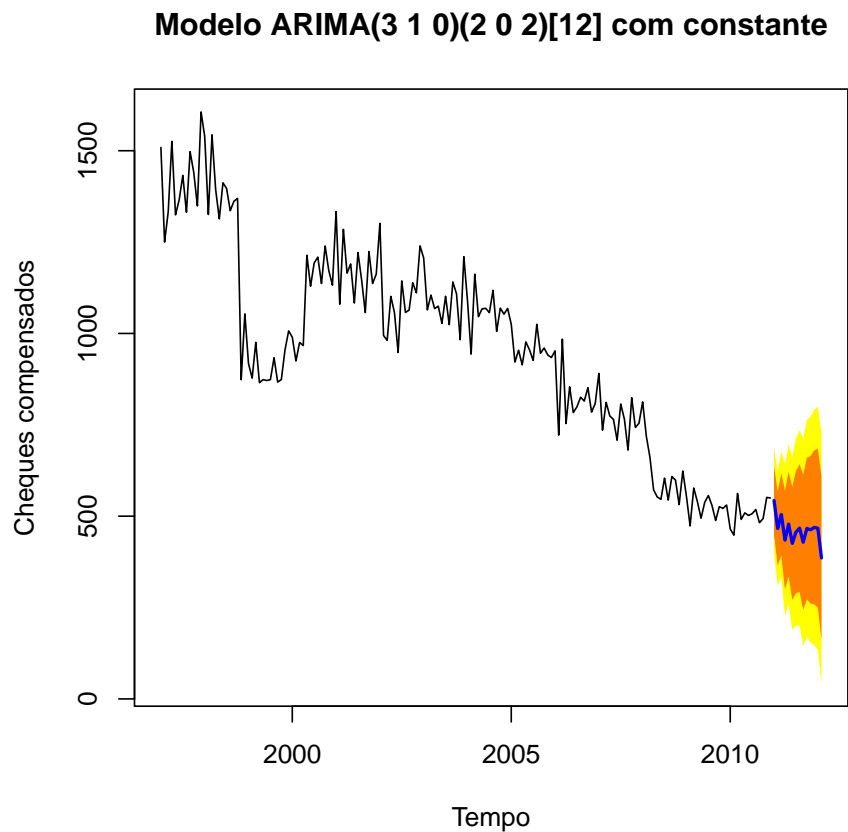


Figura 9: Previsão para quatorze meses com intervalo de confiança.

erros seguem uma distribuição normal.

Tabela 5: Valores observados e previstos para série cheques compensados e o erro associado.

mês/ano	valor observado	predição	erro
jan 2011	488.51	543.18	-54.67
fev 2011	630.79	466.37	164.42
mar 2011	850.63	504.54	346.09
abr 2011	703.06	434.43	268.63
mai 2011	715.45	478.48	236.97
jun 2011	421.12	425.59	-4.47
jul 2011	426.26	456.33	-30.07
ago 2011	428.1	467.18	-39.08
set 2011	327.97	428.79	-100.82
out 2011	352.97	466.44	-113.47
nov 2011	334.9	462.6	-127.7
dez 2011	328.09	469.3	-141.21
jan 2012	356.27	467.13	-110.86
fev 2012	325.22	385.63	-60.41

Constatou-se pelo teste Shapiro-Wilk que a hipótese de normalidade dos resíduos ficou comprometida, uma vez que o valor $P = 0.00932$, e neste ponto cabe salientar que apesar de M4 aparentemente não apresentar normalidade o modelo M12 também não apresentou pois o valor P para o teste de Shapiro-wilk sobre o erro de predição foi 0.001, e optou-se por referendar o modelo M4, pois, os demais modelos ajustados também não tiveram normalidade, e entre eles então, o eleito foi o que teve melhor adequação.

5 Conclusão

O estudo da série de cheques compensados, destinou-se a fazer o ajuste de um modelo que melhor se adeque aos dados, teve como objetivo, auxiliar os interessados, no setor econômico, onde é de grande importância prever o que irá acontecer no futuro, para determinar diretrizes, visando posteriores decisões no que se refere ao comportamento financeiro.

A análise foi feita com aplicação em séries temporais. Preliminarmente utilizou-se de processo iterativo `auto.arima` disponível no software R, que gerou um modelo com base no critério BIC. No entanto esse modelo não foi satisfatório por não apresentar ruído branco nos resíduos. Partiu-se, para eleição de forma empírica, em que foram selecionados modelos através da visualização dos correlogramas de autocorrelação e autocorrelação parcial. Os modelos sugeridos foram comparados e o eleito se deu com base na medida de precisão MAE, já que o principal objetivo do trabalho é de previsão. O modelo ajustado foi o Sarima multiplicativo de ordem $(3,1,0)(2,0,2)$ com sazonalidade de 12 meses e com uso de constante.

Referências

- BANCO CENTRAL DO BRASIL. Abril 2012. Acesso em: 30 de Abril de 2012. Disponível em: <<http://www.bcb.gov.br/pt-br/paginas/default.aspx>>.
- BEZERRA, M. I. S. *Análise de séries temporais*. São Paulo: [s.n.], 2006. Disponível em: <<http://pt.scribd.com/doc/55232564/Apostila-Series-Temporais>>.
- CARVALHO, J. J. M. de. *Modelos de previsão para cheques compensados no Brasil*. Dissertação (Mestrado) — Universidade Federal do Ceará, Fortaleza-Ceará, 2007.
- EHLERS, R. S. *Análise de Séries Temporais*. 2007. Departamento de Estatística, UFPR. Acesso em: 09 set. 2012. Disponível em: <<http://www.each.usp.br/rvicente/AnaliseDeSeriesTemporais.pdf>>.
- FISCHER, S. *Séries univariantes de tempo - metodologia de Box & Jenkins*. Porto Alegre: Fundação de Economia Estatística, 1982.
- FREITAS, N. *História do dinheiro*. Dezembro 2004. Acesso em: 15 Set. 2012. Disponível em: <<http://www.newton.freitas.nom.br/artigos.asp?cod=101>>.
- FRITSCH, W. *A Riqueza das Nações*. São Paulo: Nova Cultural, 1996. 471 p.
- HYNDMAN, R. J.; KOEHLER, A. B. Another look at measures of forecast accuracy. *Internacional Journal of Forecasting*, v. 22, p. 18, Novembro 2005.
- LOPES, J. C.; ROSSETTI, J. P. *Economia Monetária*. São Paulo: Atlas, 1992.
- MAMEDE, G. *Simulação na emissão de cheques*. São Paulo: [s.n.], 2003. Disponível em: <www.mundojuridico.adv.br/cgi-bin/upload/texto255.rtf>.
- MARTINS, F. *Títulos de Crédito*. Rio de Janeiro: Forense, 2000.
- MORAIS, E. F. *Análise de Regressão versus Séries Temporais: Uma previsão para o Índice Ibovespa*. Monografia (Graduação) — Universidade Estadual da Paraíba, Campina Grande, 2010.
- MORETTIN, P. A.; TOLOI, C. M. C. *Análise de séries temporais*. São Paulo: Edgard Blücher, 2006.
- OLIVEIRA, J. A. P. de. *Títulos de Crédito*. Porto Alegre: Livraria do advogado, 1999.
- PAIVA, C. . N.; CUNHA, A. M. *Noções de Economia*. Brasília: [s.n.], 2008. Disponível em: <[PDF] Manual de Economia.pmd - Funag www.funag.gov.br/biblioteca/index.php?option=com>.
- TEIXEIRA, E. L. *A nova lei Brasileira do cheque*. São Paulo: Saraiva, 1988.

VACONCELLOS, M. A. S.; GARCIA, M. E. *Fundamentos de Economia*. São Paulo: Saraiva, 2009.