



ROBERTO ANTÔNIO DE OLIVEIRA SAMPAIO JUNIOR

**MODELAGEM MATEMÁTICA PARA A CONSCIÊNCIA FINANCEIRA E
A BOLSA DE VALORES**

Santo André, 2018



UNIVERSIDADE FEDERAL DO ABC

CENTRO DE MATEMÁTICA, COMPUTAÇÃO E COGNIÇÃO

ROBERTO ANTÔNIO DE OLIVEIRA SAMPAIO JUNIOR

**MODELAGEM MATEMÁTICA PARA A CONSCIÊNCIA FINANCEIRA E
A BOLSA DE VALORES**

Orientador: Prof. Dr. André Ricardo de Oliveira da Fonseca

Dissertação de mestrado apresentada ao Centro de
Matemática, Computação e Cognição para
obtenção do título de Mestre

ESTE EXEMPLAR CORRESPONDE A VERSÃO FINAL DA DISSERTAÇÃO
DEFENDIDA PELO ALUNO ROBERTO ANTÔNIO DE OLIVEIRA SAMPAIO JUNIOR,
E ORIENTADA PELO PROF. DR. ANDRÉ RICARDO DE OLIVEIRA DA FONSECA.

SANTO ANDRÉ, 2018

Este exemplar foi revisado e alterado em relação à versão original, de acordo com as observações levantadas pela banca no dia da defesa, sob responsabilidade única do autor e com a anuência de seu orientador.

Santo André, 04 de Maio de 2018.

Assinatura do autor: Roberto André

Assinatura do orientador: André Ferreira



MINISTÉRIO DA EDUCAÇÃO
Fundação Universidade Federal do ABC
Programa de Pós-Graduação em Mestrado Profissional em Matemática
em Rede Nacional

Avenida dos Estados, 5001 – Bairro Santa Terezinha – Santo André – SP
CEP 09210-580 · Fone: (11) 4996-0017
profmat@ufabc.edu.br

FOLHA DE ASSINATURAS

Assinaturas dos membros da Banca Examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado do candidato Roberto Antônio de Oliveira Sampaio Junior, realizada em 16 de fevereiro de 2018:

Prof.(a) Dr.(a) **André Ricardo Oliveira da Fonseca** (Universidade Federal do ABC) – Presidente

Prof.(a) Dr.(a) **Wellington Vieira Assunção** (Universidade Federal do ABC) – Membro Titular

Prof.(a) Dr.(a) **Antonio Carlos Gracias** (Fundação Educacional Inaciana Padre Sabóia de Medeiros) – Membro Titular

Prof.(a) Dr.(a) **Sinue Dayan Barbero Lodovici** (Universidade Federal do ABC) – Membro Suplente

Prof.(a) Dr.(a) **Birajara Soares Machado** (Hospital Israelita Albert Einstein) – Membro Suplente

AGRADECIMENTOS

Agradeço...

Ao Professor Orientador André Ricardo Oliveira da Fonseca por ter expandido meus horizontes em relação à aplicabilidade da Matemática Profissional.

À UFABC pela oportunidade.

Ao curso do PROFMAT

A todos os Professores, por terem mudado significativamente minha visão e entendimento da Matemática.

À CAPES pelo apoio financeiro.

RESUMO

O intuito desse trabalho é fomentar o estudo da matemática financeira com o objetivo de um impacto social, para que os alunos de baixa renda atinjam uma consciência financeira maior durante sua formação escolar e construção de sua família. Esse estudo tem motivação pessoal e também é justificado pela falta de interesse dos alunos em assuntos de Álgebra, Lógica e Abstração. Através de modelos financeiros da modelagem matemática e de ferramentas computacionais, apresentados na forma de atividades para o Ensino Médio, espera-se uma conscientização maior do aluno em relação à sua liberdade financeira.

Palavras-chave: Juros Compostos, Modelo ARIMA, Modelagem Matemática, Modelagem Computacional.

ABSTRACT

The purpose of this work is to promote the study of financial mathematics with the objective of a social impact so that the students of low income achieve a greater financial consistency during their school formation and construction of their family. This study has personal motivation and is also justified by students' lack of interest in Algebra, Logic, and Abstraction. Through financial models, mathematical modeling and computational tools, presented in the form of activities for High School, it is expected that students will become more aware of their financial freedom.

Keywords: Compound Interest, ARIMA Model, Mathematical Modeling, Computational Modeling.

CONTEÚDO

Lista de Figuras	xix
Lista de Tabelas	xxiii
INTRODUÇÃO	1
2 UM POUCO SOBRE: CIVILIZAÇÃO, DINHEIRO E GARANTIAS	7
2.1 5 milhões - 3000 a.C.	7
2.2 3000 a.C. - 525 a.C.	8
3 CONCEITOS MATEMÁTICOS	13
3.1 Matemática Financeira	13
3.1.1 Razão e Proporção	13
Razão	13
Proporção	13
3.1.2 Porcentagem	14
3.1.3 Variação Percentual	14
Variações percentuais acumuladas	15
3.1.4 Juros	16
3.1.5 Valor atual de um conjunto de capitais	17
Sequência uniforme de pagamentos	18
Montante de uma sequência uniforme de depósitos	20
3.2 Probabilidade	21
3.2.1 Fundamentos da teoria da Probabilidade	21
Axiomas	22
A Relação dos Axiomas com dados experimentais	23
Notas de Terminologia	24
Experimento aleatório	24
Espaço Amostral ou Conjunto Elementar	25
Evento	26
Evento certo	26

	Evento Impossível	26
	Evento Simples	27
	Evento complementar	27
	Eventos mutuamente excludentes	27
	Chances de um Evento Aleatório	27
3.2.2	Função Probabilidade	29
	Modelo Equiprobabilístico.	29
	Demonstração	29
	Modelo Frequencial	30
3.2.3	Corolários imediato dos Axiomas, Probabilidade Condicional e O Teorema de Bayes	30
3.2.4	Independência	36
3.2.5	Covariância	37
3.2.6	Função Densidade de Probabilidade	37
3.3	Estatística	42
3.3.1	Tabela de Frequência	43
3.3.2	Medidas de Centralidade	44
	Mediana	44
	Moda	44
	Média Aritmética	45
	Média Geométrica	46
3.3.3	Medidas de Dispersão	47
	Desvio	48
	Desvio Médio	48
	Variância	49
	Desvio Padrão	51
3.3.4	Distribuição Binomial	51
	Média da Distribuição Binomial	53
	Desvio Padrão da Distribuição Binomial	55
3.3.5	Distribuição Normal	58
4	ECONOMETRIA FINANCEIRA	65
4.1	Séries Temporais	65
4.2	Definições iniciais	66
4.2.1	Retorno	66
	Agregação de Retornos	69

4.2.2	Especificações de um Processo Estocástico	71
4.3	Processos Estocásticos Estacionários	72
4.3.1	Propriedades da Função de Autocovariância	73
4.3.2	Funções de Autocorrelação e Autocorrelação Parcial:	74
4.3.3	Processos Lineares Estacionários	76
4.3.4	Processos Auto-regressivos	77
4.3.5	Processos de Médias Móveis	80
4.3.6	O Processo ARMA(p,q)	82
4.4	Processos Estocásticos Não Estacionários	88
4.4.1	Passeio Aleatório	89
	Recorrência/Transiência	89
4.4.2	Movimento Browniano	90
4.4.3	Processo ARIMA(p,d,q)	91
4.5	Construção de Modelos ARIMA	94
4.5.1	Identificação.	97
4.5.2	Estimação.	101
4.5.3	Verificação.	102
5	METODOLOGIA	105
5.1	Plano de aula	105
5.1.1	Aproveitamento dos alunos.	111
5.2	Yahoo Finance.	112
5.2.1	Abrindo o site.	113
5.2.2	Escolhendo a Ação.	114
5.2.3	Indo para os Dados Históricos.	115
5.2.4	Extraindo os Dados Históricos.	115
5.3	R Studio	116
5.3.1	O Arquivo CSV.	117
5.3.2	Abrindo as bibliotecas do R Studio.	119
5.3.3	Eliminando Alguns Dados	121
5.3.4	Plotando o gráfico original.	122
5.3.5	Diferenciando a série.	123
5.3.6	Plotando a FAC.	127
5.3.7	Plotando a FACP.	129
5.3.8	Os coeficientes do Modelo ARIMA.	130
5.3.9	Plotando o modelo e comparando com o Original	131

6	CONCLUSÃO	133
	Bibliografia	135
A	APÊNDICE A - QUESTIONÁRIO	137
B	APÊNDICE B - PROGRAMAÇÃO PYTHON DE MOVIMENTO BROWNIANO	139
C	APÊNDICE C - PROGRAMAÇÃO EM R STUDIO DE MODELO ARIMA	141

LISTA DE FIGURAS

Figura 1.0.1	Calculadora do Banco Central	2
Figura 1.0.2	Ampliação	2
Figura 1.0.3	Montante após 465 meses de depósitos regulares de R\$200,00	3
Figura 1.0.4	Resumo das Declarações por Faixas de Salário Mínimo. Fonte: http://www.fazenda.gov.br/centrais-de-conteudos/publicacoes/transparencia-fiscal/distribuicao-renda-e-riqueza/relatorio-distribuicao-da-renda-2016-05-09.pdf	3
Figura 1.0.5	Resumo das Declarações por Faixas de Salário Mínimo	4
Figura 3.1.1	Juros	15
Figura 3.1.2	Sequência de pagamentos	18
Figura 3.1.3	Sequência de depósitos.	20
Figura 3.1.4	Comparativo entre: Montante de uma sequência de depósitos com rendimentos(roxos), somatório de uma sequência de depósitos sem rendimentos(vermelhos) e Sequência uniforme de pagamentos de um crédito	22
Figura 3.2.1	Diagrama de Venn $A \cap B$ retirado de [6] pg. 14	32
Figura 3.2.2	Diagrama de Venn $A \cap B$ retirado de [6] pg. 15	33
Figura 3.2.3	Diagrama de árvore com as possibilidades de escolha de livros.	34
Figura 3.2.4	Diagrama de árvore com as possibilidades de alta e baixa de um ativo.	35
Figura 3.2.5	Gráfico da F.D.P. da quantidade de chuva	38
Figura 3.2.6	Gráfico da F.D.P. da quantidade de chuva de 4 a 6 ml	39
Figura 3.2.7	Gráfico da F.D.P. da quantidade de chuva de 4.5 a 5.5 ml	39
Figura 3.2.8	Gráfico da F.D.P. da quantidade de chuva de 4,75 a 5,25	40

Figura 3.2.9	Gráfico da F.D.P. da quantidade de chuva de 4.99 a 5.01	40
Figura 3.2.10	Gráfico da F.D.P. da quantidade de chuva de 0 a 10	41
Figura 4.2.1	PETR4 em escala normal	68
Figura 4.2.2	PETR4 em escala logarítmica	69
Figura 4.3.1	Função de autocorrelação de um ruído branco	77
Figura 4.3.2	Função de autocorrelação de um processo AR(1) $\phi = 0,8$ e $\phi = -0,8$	79
Figura 4.3.3	Função de autocorrelação de um processo AR(2) $\phi_1 = 0,5$ $\phi_2 = 0,3$ e $\phi_1 = 1$ $\phi_2 = -0,89$	80
Figura 4.3.4	Modelo1: AR(1) simulado	84
Figura 4.3.5	FAC do modelo 1	85
Figura 4.3.6	FACP do modelo 1	86
Figura 4.3.7	Modelo2: MA(1) simulado	87
Figura 4.3.8	ACF do modelo 2	88
Figura 4.3.9	FACP do modelo 2	89
Figura 4.4.1	Movimento Browniano gerado em Python	92
Figura 4.5.1	Gráfico Linx3	97
Figura 4.5.2	Gráfico Linx3 diferenciado 1 vez	98
Figura 4.5.3	Gráfico Linx3 Diferenciado em lag 2	98
Figura 4.5.4	Autocorrelação da série original	99
Figura 4.5.5	Autocorrelação da série diferenciada	99
Figura 4.5.6	FACP da série original	100
Figura 4.5.7	FACP da série diferenciada	101
Figura 4.5.8	Coefficientes de ϕ_k e θ_k	102
Figura 4.5.9	Simulação	102
Figura 4.5.10	Simulação 2	103
Figura 4.5.11	Simulação 3	103

Figura 4.5.12	Simulação 4	103
Figura 5.1.1	Esboço de uma aluna ao tentar demonstrar a fórmula de depósitos recorrentes.	112
Figura 5.1.2	Depoimento do Aluno 1	112
Figura 5.1.3	Depoimento do Aluno 2	113
Figura 5.2.1	Site do Yahoo Finance.	113
Figura 5.2.2	Campo de digitação.	114
Figura 5.2.3	Escolhendo a Ação.	114
Figura 5.2.4	Escolhendo a Ação.	115
Figura 5.2.5	Onde clicar.	116
Figura 5.2.6	Os Dados Históricos.	117
Figura 5.2.7	Alterando o período.	118
Figura 5.2.8	Download do período.	119
Figura 5.3.1	Tabela CSV.	120
Figura 5.3.2	R Studio - Tela inicial.	120
Figura 5.3.3	R Studio - Tela de comando.	121
Figura 5.3.4	R Studio - Abrindo as bibliotecas.	122
Figura 5.3.5	R Studio - Abrindo o Arquivo CSV.	123
Figura 5.3.6	R Studio - Eliminando dados do Arquivo CSV.	124
Figura 5.3.7	R Studio - Botão Re-run.	124
Figura 5.3.8	R Studio - Tabela de preços.	125
Figura 5.3.9	R Studio - Escrita do comando plot.	125
Figura 5.3.10	R Studio - O gráfico.	126
Figura 5.3.11	R Studio - Escrita do comando diff.	126
Figura 5.3.12	R Studio - O gráfico dos preços diferenciados 1 vez.	127
Figura 5.3.13	R Studio - Escrita do comando ACF.	128
Figura 5.3.14	R Studio - O gráfico da FAC.	128

Figura 5.3.15	R Studio - Escrita do comando PACF.	129
Figura 5.3.16	R Studio - O gráfico da FACP.	130
Figura 5.3.17	R Studio - Coeficientes ϕ e θ	131
Figura 5.3.18	R Studio - modelo1.	132
Figura 5.3.19	R Studio - modelo2.	132

LISTA DE TABELAS

Tabela 3.2.1	Tabela comparativa de Notas de Terminologia [7]	25
Tabela 3.2.2	Função Probabilidade	29
Tabela 3.3.1	Exemplo de Frequência absoluta e relativa	43
Tabela 3.3.2	Exemplo de Classes Frequência absoluta e relativa	43
Tabela 3.3.3	Histórico de taxas	46
Tabela 3.3.4	Tabela de desvio entre todos os valores	48
Tabela 3.3.5	Tabela de diferenças ao quadrado	49
Tabela 3.3.6	Relação entre as componentes elementares de erro e o erro resultante	60
Tabela 4.5.1	Tabela para ordens (1,d,0) e (0,d,1)	96
Tabela 4.5.2	Tabela para ordens (2,d,0) e (0,d,2)	96
Tabela 4.5.3	Tabela para ordens (1,d,1)	97

INTRODUÇÃO

Em momentos de dificuldade econômica nosso corpo e mente não funcionam tão bem quanto deveriam. Foi em um momento como este, de grande aflição, que comecei a idealizar como a minha vida seria se tivesse mais autonomia financeira, enquanto no mesmo momento o país decidia aumentar o tempo para a aposentadoria. Toda esta situação levou-me à reflexão sobre a minha capacidade de organizar-me financeiramente e cheguei à conclusão que ela era muito reduzida. Isto me intrigava: como eu podia ser professor de Matemática e não conseguia aumentar os números da minha conta? Por que alguns amigos (com o mesmo salário) ficavam mais ricos mesmo odiando matemática? Por que fico com sentimento negativo em relação ao governo quando ele aumenta a idade para a aposentadoria? Enfim, todas essas perguntas motivaram-me a escrever esse trabalho, que foi sendo construído aos poucos, ao passo que cada uma dessas perguntas iam sendo elucidadas.

Com isso, tive que repensar muitos hábitos na minha vida e onde depositava as minhas esperanças de estabilidade financeira. Percebi que quando sentia algo negativo em relação à aposentadoria era porque minha esperança de autonomia financeira estava sendo depositada ali e ao perceber isso, a primeira motivação do trabalho apareceu. Como não depender da aposentadoria? A primeira tentativa foi pela poupança. Percebi que um trabalhador do setor privado que se aposenta por idade, aos 65 anos, recebe um benefício médio de R\$1.197, 44 e um trabalhador do setor privado que se aposente por tempo de contribuição, ou seja, 30 anos para mulheres e 35 anos para homens, recebe um benefício médio de R\$2.303, 80, de acordo com o Boletim Estatístico de Previdência Social. Considerando uma aposentadoria de R\$2.500, 00, para que a poupança me dê esse rendimento eu teria que ter R\$500.000,00. É um valor alto, mas teria até os 60 anos para juntá-lo. Porém, a poupança está atrelada à taxa SELIC que vem caindo ao longo dos últimos anos, com isso as regras para a poupança mudariam e o dinheiro renderia menos que 0, 5%, o que faria eu ter que aumentar o montante. Mediante esta frustração fui procurar outra fonte de rendimentos e descobri: LCI, LCA, Tesouro Direto e a Bolsa de Valores.

Após a descoberta desse leque de possibilidades de investimentos percebi que manter um rendimento de 0,8% era bastante simples. Com o rendimento fixo consegui fazer algumas contas, inclusive perceber quanto tempo demoraria a me aposentar por conta própria com o mesmo valor oferecido pelo governo. Como já tenho 28 anos, fiz as contas para que me aposentasse logo. O resultado foi surpreendente, utilizando a calculadora do Banco Central concluí que em 25 anos me aposentaria com a mesma promessa do governo e eu só tinha que economizar R\$250,06 em depósitos regulares.

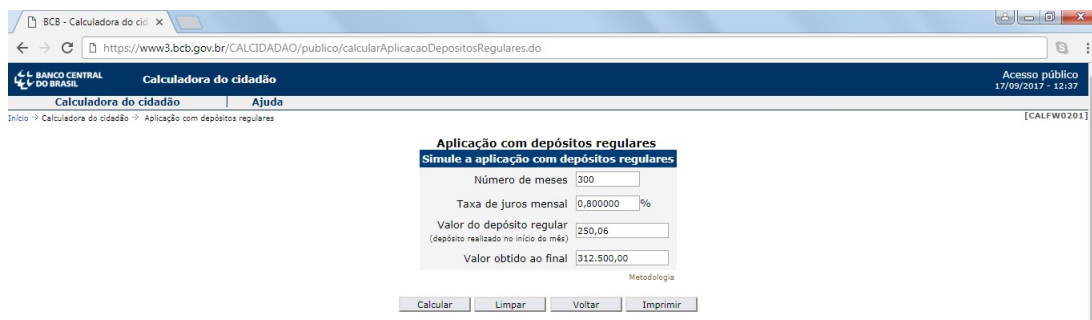


Figura 1.0.1: Calculadora do Banco Central

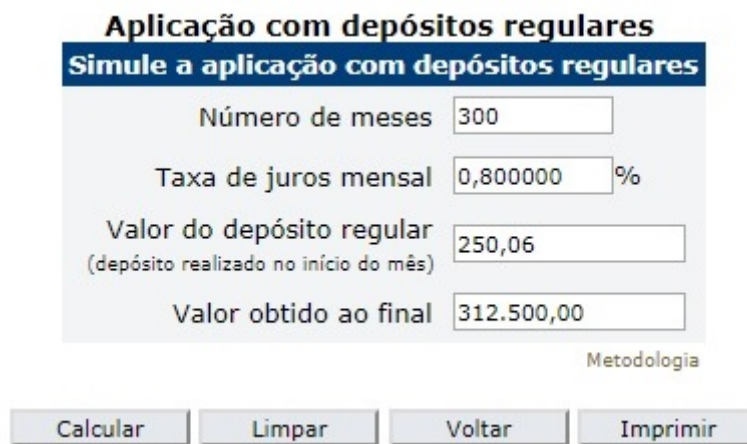


Figura 1.0.2: Ampliação

Em seguida, ao fazer outras contas percebi que para chegar a R\$1.000.000,00, era ainda mais barato, apesar de demandar mais tempo, o custo para totalizar esse montante era de apenas R\$200,00 por mês.

Ainda assim, se observarmos bem, o valor de depósito é muito inferior ao valor do montante. Veja que $200 \cdot 466 = 93.200$. Essas observações são valiosas ao ponto que

Aplicação com depósitos regulares
Simule a aplicação com depósitos regulares

Número de meses

Taxa de juros mensal %

Valor do depósito regular
(depósito realizado no início do mês)

Valor obtido ao final

Metodologia

Figura 1.0.3: Montante após 465 meses de depósitos regulares de R\$200,00

poupar R\$200,00 é acessível para muitas famílias. Veja a seguir a figura 1.0.4 de distribuição de salários no Brasil:

Valores R\$ milhões

Faixa de Rendimento	Qtde Declarantes	Rendimen to Total Bruto	Rendimen to Total Tributável	Deduçõ es Legais	Imposto Devido	Bens e Direitos Líquidos
Até 1/2 Salário Mín.	1.268.688	310	210	276	0,0	84.047
1/2 a 1 Salário Mín.	518.341	3.856	3.571	657	0,1	26.002
1 a 2 Salários Mín.	1.075.827	13.547	11.490	2.097	0,4	56.971
2 a 3 Salários Mín.	2.692.915	57.843	52.210	9.932	1,9	115.269
3 a 5 Salários Mín.	7.882.026	250.018	215.200	52.802	1.181,0	438.593
5 a 10 Salários Mín.	7.300.376	418.815	331.748	89.999	12.135,9	668.555
10 a 20 Salários Mín.	3.522.174	399.587	285.867	66.998	30.517,3	769.685
20 a 40 Salários Mín.	1.507.344	341.072	212.060	40.575	34.793,9	868.366
40 a 80 Salários Mín.	518.567	228.584	109.013	18.932	20.764,6	656.744
80 a 160 Salários Mín.	136.718	121.171	34.452	5.370	7.074,3	426.138
> 160 Salários Mín.	71.440	297.934	37.384	7.466	7.763,0	1.206.209
Total	26.494.416	2.132.738	1.293.205	295.133	114.232	5.316.579

Figura 1.0.4: Resumo das Declarações por Faixas de Salário Mínimo. Fonte: <http://www.fazenda.gov.br/centrais-de-conteudos/publicacoes/transparencia-fiscal/distribuicao-renda-e-riqueza/relatorio-distribuicao-da-renda-2016-05-09.pdf>

Tomando o salário mínimo de hoje como R\$937, 00, R\$200, 00 seriam equivalentes a 21, 34%, o que torna o investimento viável para muitos trabalhadores que não conhecem um caminho para não depender da aposentadoria oferecida pelo Governo.

Valores R\$ milhões						
Faixa de Rendimento	Qtde Declarantes	Rendimen to Total Bruto	Rendimen to Total Tributável	Deduçõ es Legais	Imposto Devido	Bens e Direitos Líquidos
Até 1/2 Salário Mín.	1.268.688	310	210	276	0,0	84.047
1/2 a 1 Salário Mín.	518.341	3.856	3.571	657	0,1	26.002
1 a 2 Salários Mín.	1.075.827	13.547	11.490	2.097	0,4	56.971
2 a 3 Salários Mín.	2.692.915	57.843	52.210	9.932	1,9	115.269
3 a 5 Salários Mín.	7.882.026	250.018	215.200	52.802	1.181,0	438.593
5 a 10 Salários Mín.	7.300.376	418.815	331.748	89.999	12.135,9	668.555
10 a 20 Salários Mín.	3.522.174	399.587	285.867	66.998	30.517,3	769.685
20 a 40 Salários Mín.	1.507.344	341.072	212.060	40.575	34.793,9	868.366
40 a 80 Salários Mín.	518.567	228.584	109.013	18.932	20.764,6	656.744
80 a 160 Salários Mín.	136.718	121.171	34.452	5.370	7.074,3	426.138
> 160 Salários Mín.	71.440	297.934	37.384	7.466	7.763,0	1.206.209
Total	26.494.416	2.132.738	1.293.205	295.133	114.232	5.316.579

Figura 1.0.5: Resumo das Declarações por Faixas de Salário Mínimo

Veja que a classe destacada na figura 1.0.5 representa 71, 52% da população declarante do Imposto, portanto ajudar essa classe a se tornar milionária no decorrer dos anos ajudaria a diminuir as diferenças sociais existentes em nosso país, além de melhorar a qualidade de vida de muitas pessoas. Pela pesquisa que fizemos (Apêndice A), observamos que:

- Aproximadamente 60% dos entrevistados não guardam dinheiro algum.
- Menos de 10% guardam mais de R\$ 400,00 por mês.

A proposta deste trabalho é despertar nos jovens, através de atividades, a motivação para guardar R\$ 200,00 todos os meses, para assim, não ficarem na dependência do governo e nem das empresas, tornando-o independente financeiramente e podendo desfrutar com segurança de uma boa terceira idade. Para isso vamos elucidar, através de atividades, como o dinheiro funciona. Para as atividades utilizaremos o livro

Fundamentos da Matemática Elementar - Vol 11- Matemática Comercial, Financeira e Estatística de Samuel Hazzan e Gelson Iezzi [10].

A Matemática Financeira é um corpo de conhecimento que estuda a variação de valor do dinheiro com o passar do tempo. Através dela, criam-se modelos que permitem avaliar e comparar o valor do dinheiro em diferentes momentos. A Matemática Financeira reconhece que dinheiro tem valor no tempo. É intuitivo entender que R\$100 em seu bolso hoje tenham mais valor do que R\$100 00 que chegarão às suas mãos daqui a seis meses [11]. Para dar contexto à sua importância, abordaremos na linha do tempo desde a idade das pedras até os dias de hoje.

Também desenvolveremos com os alunos as fórmulas utilizadas pela Calculadora do Banco Central para mostrar ainda mais a aplicação da Matemática na vida real. Discutiremos quais são as melhores formas de investimento e assim, aqueles que já trabalham podem dar início em sua aposentadoria milionária..

Conforme fui aprofundando os estudos deparei-me com o Mercado Financeiro e quis responder a seguinte pergunta em relação aos preços dos ativos:

Vai cair ou vai subir?

Para isso pude contar com a orientação do Prof. André Ricardo para entender sobre modelos AR, MA, ARMA e ARIMA com os Objetivos de:

- Modelagem;
- Previsão;
- Controle de Qualidade;
- Periodicidade.

Afim de mudar ainda mais a visão dos alunos em relação a matemática, trouxe esse assunto para debate. O assunto foi evoluindo e os alunos se interessaram mais pela bolsa de valores e perguntaram como fariam para saber se vai subir ou descer. Após esse momento selecionamos os alunos com mais aptidão e interesse para ver o funcionamento da bolsa de valores através do site “br.investing” e a utilização do modelo ARIMA na prática.

Este trabalho divide-se duas etapas interligadas. A primeira é mostrar um caminho viável, através da renda fixa, para famílias com renda até 5 salários mínimos atingirem um aporte financeiro de 1 milhão de reais, aqui considerado como suficiente para obter autonomia financeira. Poderíamos trabalhar séries do tipo P.G. de diversas formas, mas

escolhemos o caminho financeiro pelo fascínio que ele provoca e assim conseguir mais atenção dos alunos do ensino médio da escola pública para realizar os exercícios de juros retirados de [10]. No fim dos exercícios, parte dos alunos perceberam a relação entre as variáveis e conseguiram manipular as relações tempo, capital e juro. Assim caminhamos para a segunda etapa, a renda fixa. Para explorar a renda variável com os alunos, utilizamos modelos criados pelo R Studio (free software) para simular os processos estocásticos:

- AR - autorregressivos de ordem p ;
- MA - medias móveis de ordem q ;
- ARMA - junção dos dois processos anteriores;
- ARIMA - processo para séries não estacionárias, onde d é o número de vezes que a série foi diferenciada para se tornar estacionária;

E compará-los com alguns gráficos de ativos da Bolsa de valores, B3 [Brasil, Bolsa, Balcão]. Com isso, através da orientação direta em aulas teóricas e atividades, os alunos utilizaram alguns comandos do software para encontrar as variáveis (p,d,q) do processo ARIMA e realizar a simulação. O impacto social que se espera atingir com este trabalho é de que famílias de baixa renda percebam como aumentar sua renda exponencialmente usando o entendimento dos juros compostos e portanto saindo da baixa renda.

UM POUCO SOBRE: CIVILIZAÇÃO, DINHEIRO E GARANTIAS

A Matemática Financeira é uma área de conhecimento que estuda a variação de valor do dinheiro com o passar do tempo. Através dela, criam-se modelos que permitem avaliar e comparar o valor do dinheiro em diferentes momentos. A Matemática Financeira reconhece que dinheiro tem valor no tempo. É intuitivo entender que R\$100 em seu bolso hoje tenham mais valor do que R\$100, 00 que chegarão às suas mãos daqui a seis meses [11]. Para dar contexto à sua importância, abordaremos a partir da Idade da Pedra para responder perguntas: Como e quando começamos a ter a ideia de dinheiro e juros? Quando surge a ideia de acúmulo de capital? Como obter vantagens de terceiros a partir desse capital? Algumas tabuletas mostram que essas ideias não foram dos bancos mais recentes e que talvez tenha sido um acontecimento natural do pensamento humano há milhares de anos atrás. Este capítulo tenta mostrar um pouco das origens do mercado, da moeda e das estratégias usadas pelas pessoas ao longo do tempo para proteger seu capital.

2.1 5 MILHÕES - 3000 A.C.

Não se sabe quando a Idade da Pedra começou, mas há uma estimativa que seja há 5 milhões anos. Era uma época em que quase todas as pessoas eram caçadores nômades e por isso a Idade da Pedra registrou limitados avanços científicos e intelectuais. Isso não significa que faltasse inteligência às pessoas naquela época, por volta de 20 000 a.C. os caçadores das savanas haviam desenvolvido uma cultura complexa que incluía a confecção de ferramentas, linguagem, religião, arte, música e comércio. Os povos da

Idade da Pedra eram caçadores e não agricultores, portanto tinham de se deslocar de acordo com as estações. Só conseguiam levar consigo ferramentas pequenas, fáceis de transportar, roupas e objetos pessoais [5].

Antes do período Paleolítico, 10.000 a.C., o mercado de trocas já existia entre tribos asiáticas e europeias envolvendo pedras preciosas [4].

Nessa fase o homem já percebe que pode tirar proveito dos interesses de outros indivíduos, mas ainda de forma muito rudimentar, de modo que não havia leis a cerca disso, porém o suficiente para que o homem começasse a contabilizar seus pertences.

Mais adiante, entre 8000 e 5000 a.C. o mercado se expande, seguramente diferente do modo que conhecemos hoje, porém com semelhante importância. Agora o homem não vai à procura de sua comida. Na era neolítica o homem produz seu próprio alimento e assim começam a trocar sementes, ferramentas improvisadas e rebanhos de animais. Um fato interessante é que a possibilidade de acúmulo de capital através do mercado faz com que as cidades aparecessem e não o contrário [13].

Assim o capital foi mudando de forma várias vezes até o surgimento de moeda. Foi gado, grãos, joias, metais em forma pura e etc. Os pensamentos de pedir algo em troca pelo que está sendo emprestado foi aparecendo aos poucos. Gado e grão poderiam ser trocados entre os possuidores, mas também emprestados criando renda [13]. Há evidências de transições de “*Dinheiro Animal*” sendo emprestados para fazendeiros mais pobres em troca de uma parte de sua produção. Essas evidências são encontradas em diversas áreas como: Egito, América, Índia e China; e antes do aparecimento das cidades que hoje ocupam esses lugares [4].

Depois de 3000 a.C. começam a surgir as comunidades agrícolas densamente povoadas ao longo do rio Nilo na África, dos rios Tigre e Eufrates no Oriente Médio e ao longo do rio Amarelo na China. Essas comunidades criaram culturas nas quais a Ciência e a Matemática começam a desenvolver-se [5].

2.2 3000 A.C. - 525 A.C.

Enquanto antes tínhamos como classes sociais o Caçador e o Colhedor, quando o Homem aprendeu a plantar passamos a ter também o Agricultor. E ao mesmo tempo em que o pensamento Nômade ia sendo substituído, naturalmente, pela ideia dos

plantios alguns problemas agrícolas apareciam, como: períodos de secas e enchentes. Assim também apareceram as soluções: controles de irrigação, barragens e calendários.

Conforme a densidade populacional aumentava ficava mais difícil sobreviver como caçador e colhedor. Para se protegerem da fome, os povos desses lugares tiveram de encontrar outros meios de obter alimentos. Assim começaram voltar-se para a agricultura intensiva que podia alimentar populações de até 40 pessoas por milha quadrada. Isso foi uma espécie de “Revolução Agrícola” que precipitou profundas modificações culturais. Uma das mudanças foi a criação da escrita. As cheias do Amarelo, do Nilo, do Tigre e do Eufrates trouxeram a construções de barragens — atividade que requeria não só cooperação e a arte da engenharia, como também, igualmente, um sistema de preservação de registros. Os agricultores precisavam saber quando as enchentes ou a estação das chuvas chegariam, e isso significava calendários e almanaques. Os proprietários de terra mantinham anotações escritas sobre a produção agrícola e traçavam mapas que especificavam as valas de irrigação. Essas atividades deram origem a novas classes de homens educados: sacerdotes, escribas e astrólogos [5].

Outra importante mudança foi a adoção de um estilo de vida sedentário. Ao contrário dos caçadores e colhedores, os agricultores não precisavam percorrer grandes distâncias à procura de alimentos. Eles construía aldeias, vilas e pequenas cidades ao longo das margens dos rios. Perto de 2500 a.C. as cidades de Mênfis e Tebas despontavam como as metrópoles líderes do Egito. No vale do Tigre e do Eufrates, despontou primeiro a cidade de Ur, por volta de 3000 a.C. Essas primeiras cidades se agigantavam em face dos antigos lugarejos do Neolítico. Ur tinha 24 000 habitantes e uma área de 150 acres. As cidades propiciavam condições para mercados onde agricultores e artesãos podiam trocar bens, surgindo aí, para facilitar o processo, uma classe de mercadores [5].

Com o passar do tempo, mais profissões eram requeridas e assim o comércio ia ficando mais volumoso.

O comércio continuava se modificando aos poucos, agora como havia as cidades apareceria também uma espécie de volatilidade entre os produtos e por isso várias vezes intervenções estatais no comércio. Os antigos Sumérios usavam cevada na maioria das trocas. Era uma região com agricultura muito rica, assim pelos séculos houve tratados de empréstimos e pagamentos, mas ainda não havia capital em forma de moeda, então os pagamentos eram feitos em prata e grãos. As regras de pagamentos dos empréstimos variavam, assim como os valores padrões para pratas e grãos. Há muitos documentos que tratam de posses e créditos datando o tempo dos Sumérios.

Na verdade não há uma separação entre a queda dos Sumérios e o início do Império Babilônico. Além do mais, muitos costumes financeiros dos Sumérios são perpetuados no Código de Hammurabi [13].

Os Sumérios eram muito avançados nesses sentindo. Transferências eram gravadas em documentos legais e eram depositados em templos. Essa prática foi herdada dos costumes primitivos. Com o código a lei restringia a venda da propriedade principal da família e esta somente poderia ser subtraída em caso de pagamento de dívida. Uma terra não cultivada estava à disposição do primeiro ocupante que a cultivasse. O credor deveria esperar que a colheita chegasse antes de pressionar o fazendeiro pelo reembolso. Uma falha no plantio causada por seca ou tempestade anulava os juros daquele ano. As terras poderiam ser alugadas por até três anos e os pagamentos poderiam ser em produtos ou em metal precioso. [13].

As leis começam a ser estabelecidas para as profissões, incluindo o comércio. Quando um comerciante fazia trocas com tribos mais distantes ele formava uma espécie de parceria. Essas transações comerciais eram cuidadosamente reguladas pelo Código. As contas tinham que manter-se exatas; erros eram penalizados; a divisão de ganhos e provisões de capital eram detalhadas por escrito antecipadamente. Parceiros não contribuíram necessariamente com capital, eles podiam contribuir apenas com crédito e emprestar a soma necessária [13].

Assim os primórdios do que temos hoje estava formado.

Pela primeira vez na história, alguns povos tinham tempo de lazer. Enquanto os agricultores, que formavam a maioria da população durante a Revolução Agrícola, gastavam o dia todo no trabalho, outras pessoas — reis, sacerdotes, mercadores e escribas — tinham tempo ao fim do dia para ponderar sobre os mistérios da natureza e da ciência. Por fim, todos os ingredientes para o progresso científico estavam reunidos: escrita, necessidade de novas tecnologias, ambientes urbanos e tempo de lazer [5].

E de acordo com [5], as tribos foram sendo substituídas por Cidades-Estados e o chefe tribal por vastas burocracias que nos acompanham até hoje. Podemos observar que a ideia de retorno e proteção do capital investido é tão antigo quanto à própria civilização, daí importância de estudar Matemática Financeira e suas aplicações.

Não nos cabe aqui estudar a evolução do dinheiro em si, mas os meios que foram criados para proteção de capital acumulado, assim como ganhos sobre ele. Uma ciência que começou há tanto tempo, obviamente sofreu muitas evoluções e ramificações. Hoje em dia, temos o mercado financeiro e seus órgãos regulamentadores, cada país

com o seu, assim como sempre foi, desde os mercados babilônicos, sumérios e outros dessa época antiga. Daremos aqui ênfase ao mercado brasileiro, a BVM&F.

Na prática, esta parte do trabalho serviu como contextualização do tema através da história, o que traz mais motivação aos alunos.

3

CONCEITOS MATEMÁTICOS

Alguns conhecimentos prévios são necessários para o domínio das habilidades de trabalho na bolsa de valores. A seguir vamos apresentar as definições desses conhecimentos em 3 seções. A primeira sobre temas da matemática financeira, a segunda em eventos probabilísticos e a terceira discutiremos alguns aspectos estatísticos.

3.1 MATEMÁTICA FINANCEIRA

3.1.1 *Razão e Proporção*

Razão

Dado dois números a e b , com $b \neq 0$, chama-se de razão entre a e b o quociente $\frac{a}{b}$.

Exemplo: Há um total de 30 frutas: 15 bananas, 10 abacaxis e 5 maçãs. A razão entre o número de maçãs e o número de bananas é dado por $\frac{5}{15}$, simplificando obtemos $\frac{1}{3}$ significando que temos uma maçã à cada 3 bananas. Ou ainda temos a razão de bananas para o total de frutas $\frac{15}{30}$, simplificando obtemos $\frac{1}{2}$ significando que bananas são metade do total de frutas.

Proporção

A igualdade entre duas razões é denominada proporção. Exemplo: Na proporção $\frac{5}{15} = \frac{1}{3}$ dizemos que 5 está para 15 assim como 1 está para 3.

A propriedade fundamental da proporção diz que o produto dos meios deve ser igual ao produto dos extremos. Exemplo: $5 \div 15 = 1 \div 3 \rightarrow 1 \cdot 15 = 5 \cdot 3$

Essa propriedade nos permite resolver alguns tipos de problemas, veja:

Um médico precisa aplicar 2ml de um remédio para cada kg de uma pessoa. Se a pessoa pesa 75 kg, quantos ml de medicamento são necessários?

$\frac{2\text{ml}}{1\text{kg}} = \frac{x\text{ml}}{75\text{kg}}$ como o produto dos meios deve ser igual ao dos extremos, temos:

$1 \cdot x = 2 \cdot 75 \rightarrow x = 150$, ou seja, o médico deve aplicar uma quantidade de 150 ml de remédio.

3.1.2 Porcentagem

Dado um número a qualquer, ao escrever $a\%$ - lê-se 'a' por cento - estamos representando a razão $\frac{a}{100}$. Por cento é uma expressão representada pelo símbolo %, que significa centésimos. Logo $12\% = \frac{12}{100}$ Exemplos:

- a) $25\% = \frac{25}{100} = 0,25$
- b) $3,8\% = \frac{3,8}{100} = 0,038$
- c) $145\% = \frac{145}{100} = 1,45$
- d) $401,7\% = \frac{401,7}{100} = 4,017$

3.1.3 Variação Percentual

Suponha que uma peça de carro custe R\$80 reais. Por razões comerciais esse produto passa a custar R\$84 no mês seguinte. A razão $\frac{a}{b}$, a = aumento e b = valor inicial, é chamada **variação percentual** entre as datas consideradas.

$$j = \frac{4}{80} = 0,05 = 5\%$$

De modo geral, consideremos uma grandeza de valor V_0 na data inicial e o valor V_f na data final. Podemos escrever a variação percentual da seguinte forma

$$j = \frac{V_f - V_0}{V_0} \quad (3.1.1)$$

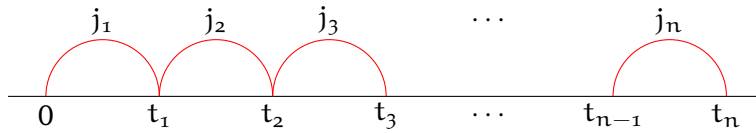


Figura 3.1.1: Juros

Sendo j igual ao número obtido como a variação percentual. Ainda podemos escrever de outra forma o mesmo procedimento

$$j = \frac{V_f}{V_o} - 1 \quad (3.1.2)$$

Observemos que quando $V_f > V_o$ temos uma variação positiva e assim uma taxa percentual de crescimento. Quando ocorre o contrário temos uma taxa percentual de decrescimento.

Variações percentuais acumuladas

Considerando os instantes de tempo $0, t_1, t_2, t_3, \dots, t_{n-1}, t_n$, em que $t_{k-1} < t_k$ e ainda chamemos j_k a variação percentual entre os instantes t_{k-1} e t_k ; cada j_k é uma variação sucessiva de j_{k-1}

Dado uma grandeza inicial em $t=0$, podemos expressar as seguintes informações:

$$j_1 = \frac{V_1}{V_0} - 1 \Rightarrow V_1 = V_0(1 + j_1)$$

$$j_2 = \frac{V_2}{V_1} - 1 \Rightarrow V_2 = V_1(1 + j_2) \Rightarrow V_2 = V_0(1 + j_1)(1 + j_2)$$

$$j_3 = \frac{V_3}{V_2} - 1 \Rightarrow V_3 = V_2(1 + j_3) \Rightarrow V_3 = V_0(1 + j_1)(1 + j_2)(1 + j_3)$$

Assim podemos observar que:

$$V_n = V_0(1 + j_1)(1 + j_2)(1 + j_3) \cdots (1 + j_n) \quad (3.1.3)$$

Para encontrar a variação percentual acumulada (j_{ac}) devemos aplicar a equação 3.1.2, logo:

$$j_{ac} = \frac{V_n}{V_o} - 1$$

Substituindo o numerador, temos:

$$j_{ac} = \frac{V_0(1+j_1)(1+j_2)(1+j_3)\cdots(1+j_n)}{V_0} - 1$$

$$j_{ac} = (1+j_1)(1+j_2)(1+j_3)\cdots(1+j_n) - 1 \quad (3.1.4)$$

Vale destacar que ao calcular as **taxas de variações acumuladas**, em termo real, estaríamos calculando as taxas de inflação/deflação de um determinado período.

3.1.4 Juros

Existem diversas formas para definir juros. Para este trabalho vamos definir juros como a taxa de retorno, positiva ou negativa, de um investimento em determinado tempo. Onde J é o juro, M o montante acumulado após um período de tempo e C o capital inicial do investimento. No mundo financeiro não há uma aplicação que se comporte como juros simples, portanto consideraremos apenas o juro composto. Assim dado um capital inicial C , uma taxa i de juro e um tempo n temos:

$$M_n = C \cdot (1+i)^n \quad (3.1.5)$$

Exemplo:

Foi aplicado um capital de R\$3000,00 em um investimento que possui uma taxa de retorno a 2% ao mês.

- O montante ao final de 3 meses é dado por:

$$M_3 = 3000 \cdot (1 + 0,02)^3 \rightarrow M_3 = 3000 \cdot 1,061208 \rightarrow M_3 = 3183,64$$

- Ou ainda, no final de 6 meses:

$$M_6 = 3000 \cdot (1 + 0,02)^6 \rightarrow M_6 = 3000 \cdot 1,126162419264 \rightarrow M_6 = 3378,49$$

Observe que $3378,49 - 3183,64 = 194,84$, ou seja, nos 3 últimos meses a rentabilidade foi maior que nos 3 primeiros meses, isso ocorre devido ao juro composto, que é juro sobre juro. O fator

$$(1+i)^n \quad (3.1.6)$$

É chamado de fator de ágio, que é a correção temporal, no futuro, para o capital inicial após o período n . Todavia podemos encontrar o capital inicial a partir do montante e da data futura, para isso temos o fator de deságio que é o inverso do fator de ágio:

$$\frac{1}{(1+i)^n} = (1+i)^{-n} \quad (3.1.7)$$

Enquanto 3.1.6 nos permite ir do presente ao futuro de um investimento, a 3.1.7 nos permite vir do futuro para o presente. Juros então significa a taxa de atualização temporal do valor do dinheiro, uma variação determinada quando o dinheiro é transportado pelo tempo.

3.1.5 Valor atual de um conjunto de capitais

Suponhamos um devedor e um credor de uma dívida de R\$10.000,00 que vença daqui um mês. Se o credor conhece uma aplicação à taxa de 0,8%, qual seria o valor justo para o devedor quitar a dívida antecipadamente? Para isso vamos aplicar a equação 3.1.7.

$$C = \frac{10000}{1,008} = 9.920,63$$

Ou seja, um valor aceitável pelo credor seria R\$9.920,63.

Suponha ainda que esse mesmo credor tenha que receber outra dívida de R\$16.000,00 daqui 2 meses e seu devedor lhe oferece R\$15.500,00 a vista, o credor deve aceitar a proposta?

$$C = \frac{16000}{(1,008)^2} = 15.747,03$$

Logo vemos que não é interessante para o credor receber a dívida hoje no valor proposto pelo devedor.

Vamos resolver agora a seguinte situação: uma loja vende um fogão a prazo em 5 prestações mensais de R\$400,00, sendo a primeira paga após um mês à compra. Se o pagamento for à vista, o preço cobrado é R\$1850,00. Conhecendo uma aplicação que renda 0,8%, é mais vantajoso para o comprador a opção à vista ou à prazo? Para responder essa questão o comprador precisa trazer para o presente cada uma das parcelas de R\$400,00:

$$Vp = \frac{400}{1,008} + \frac{400}{(1,008)^2} + \frac{400}{(1,008)^3} + \frac{400}{(1,008)^4} + \frac{400}{(1,008)^5}$$

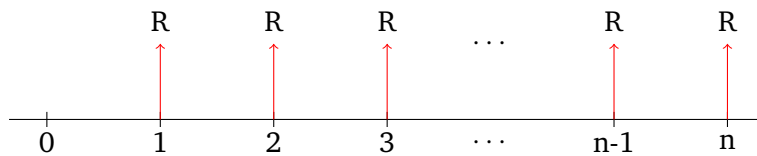


Figura 3.1.2: Sequência de pagamentos

$$Vp = 396,82 + 393,68 + 390,55 + 387,45 + 384,38 = 1952,88$$

Com o cálculo o comprador percebe que a compra a vista é melhor opção. Nessa mesma loja uma televisão é vendida a vista por R\$1.100,00 ou em duas prestações mensais a um juro composto de 3%. Qual deve ser o valor de cada parcela? Para calcular o valor das parcelas devemos levar o valor 1 mês à frente

$$Vf_1 = 1100 \cdot 1,03 = 1133$$

como serão duas parcelas, vamos dividir o valor em 2 e descontar do valor Vf_1

$$Vp_1 = 1133 - \frac{1133}{2} = 566,5$$

agora devemos levar esse valor mais um mês à frente

$$Vf_2 = Vp_2 = 566,5 \cdot 1,03 = 583,50$$

Assim o valor de cada prestação deve ser a soma dos valores parciais divididos por 2

$$V_{prest} = \frac{Vp_1 + Vp_2}{2} = \frac{566,5 + 583,5}{2} = 575$$

o valor de cada parcela, então, é R\$575,00.

Sequência uniforme de pagamentos

Considere um valor V que deve ser financiado, por exemplo, R\$25.000,00 para a compra de um carro. Se o banco empresta esse dinheiro a uma taxa de 3% ao mês, qual o valor de cada prestação se a dívida deve ser liquidada em 2 anos. Chamaremos esse conjunto de pagamentos de **sequência uniforme de pagamentos**. Veja na figura a seguir como o algoritmo funciona de modo geral:

Podemos escrever o valor V atual das n prestações, à taxa j como:

$$V = \frac{R}{(1+j)} + \frac{R}{(1+j)^2} + \frac{R}{(1+j)^3} \cdots + \frac{R}{(1+j)^n}$$

Podemos observar que o valor de V é uma Progressão Geométrica de razão $q = \frac{1}{(1+j)}$, cujo primeiro termo é $\frac{R}{(1+j)}$, assim podemos aplicar a soma de uma P.G. finita

$$S_n = a_1 + a_2 + a_3 + \cdots + a_n = \frac{a_1(q^n - 1)}{(q - 1)} \quad (3.1.8)$$

Substituindo os valores, temos:

$$\begin{aligned} V &= \frac{\frac{R}{(1+j)} \left[\frac{1}{(1+j)^n} - 1 \right]}{\frac{1}{(1+j)} - 1} \\ V &= R \cdot \frac{\frac{1}{(1+j)} \left[\frac{1-(1+j)^n}{(1+j)^n} \right]}{\frac{1-(1+j)}{(1+j)}} \\ V &= R \cdot \frac{\left[\frac{1-(1+j)^n}{(1+j)^n} \right]}{-j} = R \cdot \frac{\left[\frac{(1+j)^n - 1}{(1+j)^n} \right]}{j} \end{aligned}$$

e por fim:

$$V = R \cdot \frac{(1+j)^n - 1}{(1+j)^n \cdot j} \quad (3.1.9)$$

Essa é a fórmula que relaciona o valor atual com a prestação de valor R , taxa de juros j e o número n de prestações. Assim o cálculo para o nosso problema seria:

$$\begin{aligned} 25.000 &= R \cdot \frac{(1,03)^{24} - 1}{(1,03)^{24} \cdot 0,03} \\ 25.000 &= R \cdot \frac{1,032}{0,061} \\ 1477,32 &= R \end{aligned}$$

Conclui-se que R\$1477,32 é o valor de cada uma das 24 parcelas. Supondo agora outra situação parecida, vamos imaginar que nosso tomador de crédito não queira assumir essa dívida por 2 anos, mas por apenas um e que consiga pagar apenas R\$1200,00 por mês, quanto o gerente pode emprestar a ele hoje? Para essa resposta o cálculo é o seguinte:

$$\begin{aligned} V &= 1200 \cdot \frac{(1,03)^{12} - 1}{(1,03)^{12} \cdot 0,03} \\ V &= 1200 \cdot \frac{0,425}{0,0425} \end{aligned}$$

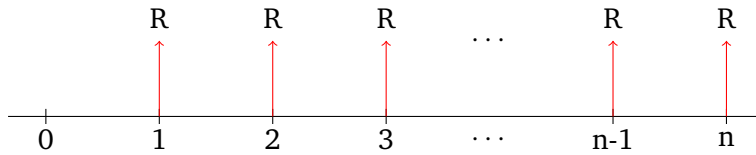


Figura 3.1.3: Sequência de depósitos.

$$V = 12.000$$

Ou seja, o valor que o gerente pode emprestar ao nosso tomador de crédito é R\$12.000,00. É fácil verificar que $1.200 \cdot 12 = 14.400$ um montante maior do que o valor oferecido pelo banco. Isso se deve ao deságio do dinheiro quando o trazemos para o presente. Emprestar dinheiro, então, é receber um valor menor do que o valor pago.

Montante de uma sequência uniforme de depósitos

Vamos agora analisar um poupador, situação totalmente adversa à anterior. Um jovem inicia uma sequência de depósitos de R\$200,00 com a intenção de realizar uma viagem no fim da faculdade. Por fim, passam-se os 72 meses de curso. Ele esperava um saldo de R\$14.400 e para a sua surpresa, ao tirar o extrato da conta, o valor obtido era R\$19.525,88. Vejamos o que ocorreu:

Consideremos uma sequência de depósitos iguais de valor R , nas datas $1, 2, 3, \dots, n$, rendendo a uma taxa de juros j , se quisermos saber o montante ao final devemos utilizar a equação 3.1.6 em cada depósito na data $k \leq n$.

- montante do 1º depósito na data n : $R \cdot (1 + j)^{n-1}$
- montante do 2º depósito na data n : $R \cdot (1 + j)^{n-2}$
- montante do 3º depósito na data n : $R \cdot (1 + j)^{n-3} \dots$
- montante do n° depósito na data n : $R \cdot (1 + j)^{n-n} = R$

Assim:

$$M_t = R \cdot (1+j)^{n-1} + R \cdot (1+j)^{n-2} + R \cdot (1+j)^{n-3} + \dots + R$$

Os termos, assim como na sequência uniforme de pagamentos, expressam uma Progressão Geométrica de razão $q = \frac{1}{1+j}$ e cujo 1º termo é $a_1 = R(1+j)^{n-1}$. Ao aplicar a fórmula da soma de P.G. a equação 3.1.8:

$$M = \frac{R \cdot (1+j)^{n-1} \left(\frac{1}{(1+j)^n} - 1 \right)}{\frac{1}{1+j} - 1}$$

$$M = R \cdot \frac{\frac{1}{1+j} - (1+j)^{n-1}}{\frac{-j}{1+j}}$$

$$M = R \cdot \frac{\frac{1-(1+j)^n}{1+j}}{\frac{-j}{1+j}}$$

$$M = R \cdot \frac{1 - (1+j)^n}{-j}$$

e por fim:

$$M = R \cdot \frac{(1+j)^n - 1}{j} \quad (3.1.10)$$

Podemos observar no gráfico a seguir, a história monetária de quem pede um valor emprestado ao banco a uma taxa de 3% e de quem guarda á uma taxa de 0,6%. Por questão de simplicidade, os gráficos trabalham com depósitos/pagamentos regulares de R\$1,00, assim qualquer conversão pode ser feita com simplicidade. A linha vermelha significa a somatória dos depósitos/pagamentos regulares se não tivessem sido investidos ou sendo pagos ao banco.

3.2 PROBABILIDADE

3.2.1 Fundamentos da teoria da Probabilidade

Vamos trazer algumas definições que estão presentes nos trabalhos de [7] e [6], além de serem trabalhos que fundamentam muito bem as bases da Probabilidade, pois a apresentam com axiomas e definições, assim como faz a Geometria e a Álgebra,

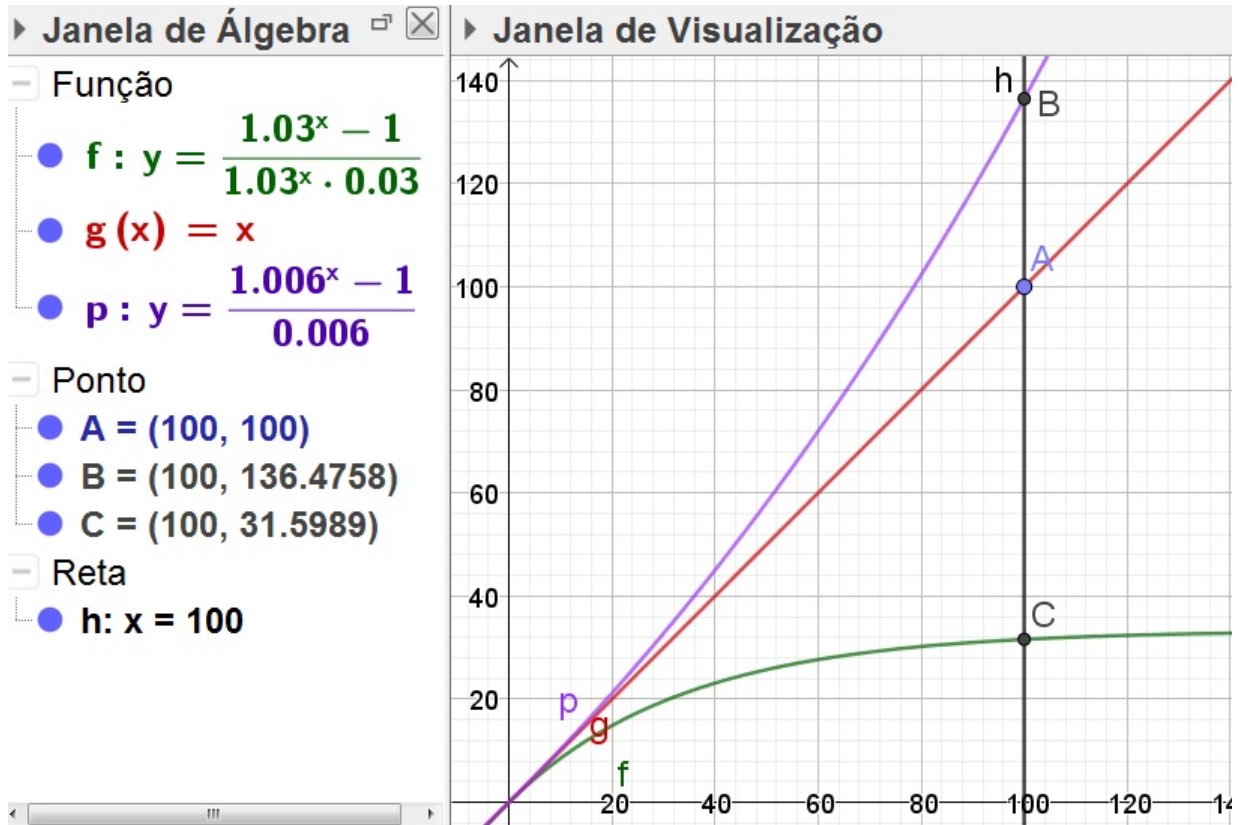


Figura 3.1.4: Comparativo entre: Montante de uma sequência de depósitos com rendimentos(roxo), somatório de uma sequência de depósitos sem rendimentos(vermelho) e Sequência uniforme de pagamentos de um crédito

são também trabalhos em que os autores se preocuparam em desenvolver a Teoria da Probabilidade de forma próxima à probabilidade empírica.

Axiomas

Em [7],

Afirmativa 1. *Seja E uma coleção de elementos $\varepsilon, \eta, \zeta, \dots$, que serão chamados de eventos elementares, e F um conjunto de subconjuntos de Ω ; os elementos de F serão chamados de eventos aleatórios.*

I- F é um campo/classe de conjuntos de eventos aleatórios

II- F contém Ω

III- Para cada A em F é associado um número real não negativo $P(A)$. O número $P(A)$ é chamado de probabilidade do evento A .

IV- $P(\Omega) = 1$

V- Se A e B não tem elementos em comum, então

$$P(A + B) = P(A) + P(B)$$

Se o sistema de conjuntos satisfaz os axiomas I-V, ele é chamado de *campo de probabilidade*.

Esse sistema de axiomas é *consistente* e podemos verificar pelo seguinte exemplo, tirado de [7]. Seja Ω um conjunto unitário como elemento ε e o conjunto F contendo Ω e o conjunto nulo \emptyset . $P(\Omega) = 1$ e $P(\emptyset) = 0$. Agora para a construção do campo de probabilidade podemos pegar, arbitrariamente, um conjunto finito $\Omega = \{\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_k\}$ e um conjunto, também arbitrário, $\{p_1, p_2, \dots, p_k\}$ de números não negativos, com soma $p_1 + p_2 + \dots + p_k = 1$. Tomamos F como conjunto de todos os subconjuntos em Ω e escrevemos

$$P\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_k\} = \{p_1, p_2, \dots, p_k\}$$

nesse caso cada p_i é uma probabilidade de um evento elementar.

A Relação dos Axiomas com dados experimentais

Kolmogorov [7], mostra como a teoria da probabilidade pode se aplicar a situações do mundo real:

1. Assumimos um complexo de condições quaisquer, chamado “ F ”, que permite, de forma arbitrária, um número qualquer de repetições.
2. Dessas condições vamos estudar um conjunto “ Ω ” de acontecimentos possíveis dentro dessas condições. Para determinados conjuntos alguns acontecimentos não são coerentes, então escolheremos para o conjunto Ω apenas aqueles com chances de ocorrer.
3. Se um evento ocorre dentro das condições definidas e o conjunto de acontecimentos possíveis, dizemos que o evento “ a ” ocorreu.

Exemplo tirado de [7], considerando o lançamento de duas moedas com resultados possíveis cara ou coroa. Foi estabelecido aqui um complexo de condições e

agora vamos determinar o conjunto E de possíveis acontecimentos sendo cara K e coroa C ,

$$E = \{CC, CK, KC, KK\}$$

O evento “ a ”, necessariamente, é um desses quatro *eventos elementares* de E , quando o evento a qualquer não for um desses, então dizemos que a não pode ocorrer.

4. Podemos associar o evento a ao número real $P(a)$, desde que o complexo de condições tenha as seguintes características:
 - a) Temos certeza que o complexo de condições possa ser repetido quantas vezes quisermos, por exemplo um número muito grande n , então se m é o número de vezes que ocorreu o evento a , a razão entre m e n será muito semelhante ao $P(A)$.
 - b) Se $P(A)$ é muito pequeno e se as condições F acontecerem apenas uma vez, podemos ter certeza que o evento A não vai ocorrer de forma alguma.

Notas de Terminologia

Em seu trabalho, [7], monta uma relação dos eventos aleatórios com a Teoria dos Conjuntos, podemos observá-la na tabela 3.2.1:

Experimento aleatório

De acordo com [7], é todo fenômeno que apresenta resultado imprevisível, mesmo quando repetido inúmeras vezes nas mesmas circunstâncias. Exemplo:

1. Uma pessoa pensa em um número de 1 a 10;
2. Um sorteio das bolas de bingo;
3. Retirar uma carta do baralho;
4. Lançar um dado ou uma moeda.

Tabela 3.2.1: Tabela comparativa de Notas de Terminologia [7]

Eventos Aleatórios	Teoria dos Conjuntos
1. Evento A e B são incompatíveis.	1. A e B tem intersecção vazia. Em notação $AB=0$
2. Eventos A,B,C...N são incompatíveis	2. $ABC...N=0$
3. O evento X é definido como a ocorrência simultânea dos eventos A,B,C..N	3. $ABC...N=X$
4. O evento X e definido como a ocorrência de pelo menos um dos eventos A,B,C...N	4. $A+B+C+...+N=X$
5. O oposto de A consiste na não ocorrência do mesmo.	5. É o conjunto A^c
6. A é um evento impossível.	6. $A=\emptyset$
7. O evento A tem que ocorrer.	7. $A=E$
8. Um experimento κ consiste em determinar quais dos eventos A_1, A_2, \dots, A_n ocorrem. Podemos chamar A_1, A_2, \dots, A_n de possíveis resultados de κ	8. Um sistema κ de conjuntos A_1, A_2, \dots, A_n formam a decomposição do conjunto E, se $A_1 + A_2 + \dots + A_n = E$. (Assumindo que cada A_x são disjuntos dois a dois.)
9. A ocorrência do evento B segue inevitavelmente da ocorrência de A.	9. B é subconjunto de A: $B \subset A$

Espaço Amostral ou Conjunto Elementar

O conjunto de todos os possíveis acontecimentos de um experimento aleatório. Em geral esse conjunto é denominado pela letra grega ômega - Ω . Exemplos

1. Os 10 possíveis números que uma pessoa pode imaginar. $\Omega = \{1, 2, 3, \dots, 9, 10\}$, onde o número de elementos é $n(\Omega) = 10$
2. Todas as bolas do bingo. $\Omega = \{1, 2, 3, \dots, 74, 75\}$, onde o número de elementos é $n(\Omega) = 75$
3. Ao lançar duas vezes uma moeda e observar a face voltada para cima, sendo elas cara (k) e coroa (c). $\Omega = \{(k, k)(k, c)(c, k)(c, c)\}$, onde o número de elementos é $n(\Omega) = 4$ (usado também em [7])

No trabalho de [6] encontramos a seguinte definição

Definição 1. Seja Ω o espaço amostral do experimento. Todo subconjunto $A \subset \Omega$ será chamado *evento*.

Evento

Em [7] encontramos que é qualquer subconjunto de um espaço amostral Ω como acabamos de definir. Em contexto seria um fato esperado. Representaremos por E_i . Exemplos:

1. No lançamento de 2 moedas podemos esperar o evento E_1 de aparecerem faces iguais.
 $E_1 = \{(k, k)(c, c)\}$ Portanto o número de elementos do evento E_1 é $n(E_1) = 2$
 Ou ainda um evento E_2 de aparecer coroa em pelo menos um lançamento.
 $E_2 = \{(c, k), (k, c), (c, c)\}$ onde $n(E_2) = 3$
2. No lançamento de um dado de seis faces.
 E_1 : o resultado ser par. $E_1 = \{2, 4, 6\}$ onde $n(E_1) = 3$

Agora alguns casos particulares (Alguns já relacionados pela tabela 3.2.1 de [7]):

Evento certo é aquele que possui os mesmo elementos do espaço amostral, ou seja,

$$E = \Omega$$

Evento impossível não possui elemento comum com o espaço amostral, isto é,

$$E \cap \Omega = \emptyset$$

Evento Simples possui um único elemento, $n(E) = 1$

Evento complementar se E é um evento em um espaço amostral Ω , seu evento complementar é indicado por E^c e tal que $E^c = \Omega \setminus E$

Eventos mutuamente excludentes são dois ou mais eventos que a ocorrência de um deles implica na não ocorrência do outro. Se E e D são eventos mutuamente excludentes, então $E \cap D = \emptyset$

No trabalho de [6]:

Definição 2. Um evento A ao qual atribuímos uma probabilidade será chamado de *evento aleatório*

Chances de um Evento Aleatório

Em seu trabalho, Barry [6] nos fornece algumas proposições antes de definir bem as *Chances de um Evento Aleatório*:

Propriedade 1. $\Omega \in F$ (Definiremos $P(\Omega) = 1$)

Propriedade 2. Se $A \in F$, então $A^c \in F$. É evidente que

$$P(A^c) = 1 - P(A)$$

Propriedade 3. Se $A \in F$ e $B \in F$, então $A \cup B \in F$

Definição 3. Se [

a Ω um conjunto não-vazio. Uma classe F de subconjuntos de Ω satisfazendo as propriedades 1,2,3 é chamada álgebra de subconjuntos de Ω .

Proposição 1. Seja F uma álgebra de subconjuntos de Ω , então valem as seguintes propriedades:

Propriedade 4. $\emptyset \in F$

Propriedade 5. $\forall n \forall A_1, A_2, \dots, A_n \in F$, temos

$$\bigcup_{i=1}^n A_i \in F$$

e

$$\bigcap_{i=1}^n A_i \in F$$

A proposição 1 diz que uma álgebra é fechada para um número finito de aplicações das operações \cup , \cap e c

Demonstração. As propriedades 1 e 2 implicam na propriedade 4. Para mostrar a propriedade 5, podemos utilizar a propriedade 3 e temos que

$$A_1 \cup A_2 \in F$$

assim

$$(A_1 \cup A_2) \cup A_3 \in F$$

e por indução temos que

$$\bigcup_{i=1}^n A_i \in F$$

Agora basta observar que

$$\bigcap_{i=1}^n A_i = \left(\bigcup_{i=1}^n A_i^c \right)^c$$

□

A perspectiva favorável de que algo venha a ocorrer ou as chances de tal acontecimento com um grau de segurança com que se pode esperar a realização de um evento é determinado pela *Razão* entre frequência relativa dos eventos do mesmo tipo numa série de tentativas e o total de possibilidade (*Espaço Amostral*, como o número de possibilidades de um evento esperado é menor ou igual ao total de eventos possíveis, vamos relacionar essa segurança a um número positivo entre zero e um, onde 1 é a certeza de que o evento ocorrerá e 0 é a certeza que o evento não ocorrerá, onde quaisquer valores entre 1 e 0 é a taxa de segurança de tal evento ocorrer.

$$P(E) = \frac{n(E)}{n(\Omega)} \quad (3.2.1)$$

3.2.2 Função Probabilidade

A função de probabilidade faz corresponder a cada valor x do espaço de resultados - que é obrigatoriamente um conjunto enumerável - a um valor y real não negativo menor ou igual a 1, valor esse que indica a probabilidade da variável aleatória DISCRETA X para o valor x .

Em outras palavras, seja Ω o espaço amostral, e $f : \Omega \rightarrow \mathbb{R}$ a função de probabilidade. Então temos que:

$$0 \leq f(x) \leq 1$$

Ou seja a imagem que define certo espaço amostral Ω deve ser o intervalo

$$P(\Omega) \rightarrow [0, 1]$$

Exemplo:

$N = \{1, 2, 3, 4, 5\}$ e $A = \{0, 1\}$ (Para A fazemos a seguinte leitura: 0 significa falso e 1 verdadeiro)

X : é número ímpar (X variável aleatória DISCRETA) $X: N \rightarrow A$

X	0	1
$f(x)$	$\frac{2}{5}$	$\frac{3}{5}$

Tabela 3.2.2: Função Probabilidade

Modelo Equiprobabilístico.

Considere um espaço amostral com t elementos. Então este espaço amostral possui t eventos unitários e a probabilidade de cada um desses eventos unitários ocorrer é $\frac{1}{k}$ para $x \in \Omega$.

Demonstração

$$\Omega = x_1, x_2, x_3, \dots, x_k \text{ e } P(x_1) = P(x_2) = P(x_3) = \dots P(x_k) = y$$

$$1 = P(\Omega) = P(x_1, x_2, x_3, \dots, x_k)$$

$$\begin{aligned}
&= P(x_1) \cup P(x_2) \cup \dots \cup P(x_k) \\
&= P(x_1) + P(x_2) + \dots + P(x_k) \\
&= y + y + y + \dots + y = k \cdot y \\
&\text{assim, } y = \frac{1}{k}
\end{aligned}$$

Modelo Frequencial

Ao repetir a experiência t vezes e o evento E ocorrer em k dessas experiências, adota-se para $P(E)$ a frequência relativa do evento E , a razão entre o número de vezes que o evento E ocorreu e o número total de repetições da experiência, assim $P(E) = \frac{k}{t}$. Formalmente o modelo frequencial ou modelo estatístico considera o limite de frequências relativas como o valor da probabilidade. portanto:

$$P(E) = \lim_{t \rightarrow \infty} \frac{k}{t} \quad (3.2.2)$$

3.2.3 Corolários imediato dos Axiomas, Probabilidade Condicional e O Teorema de Bayes

Na obra de [7] temos que $A \cup A^c = E$ e pelos axiomas (IV) e (V) segue que:

$$P(A) + P(A^c) = 1 \quad (3.2.3)$$

$$P(A) = 1 - P(A^c) \quad (3.2.4)$$

Como $E^c = \emptyset$ então

$$P(\emptyset) = 0 \quad (3.2.5)$$

Se A, B, \dots, N são excludentes, então do Axioma (V) segue a formula do *Teorema da Adição*:

$$P(A + B + \dots + N) = P(A) + P(B) + \dots + P(N) \quad (3.2.6)$$

Se $P(A) > 0$, então o quociente

$$P_A(B) = \frac{P(A \mid B)}{P(A)} \quad (3.2.7)$$

A equação 3.2.7 é definida como *Probabilidade Condicional do evento B dado que A ocorreu*.

De 3.2.7 segue, imediatamente, que:

$$P(A | B) = P(A)P_A(B) \quad (3.2.8)$$

E por indução encontra-se o *Teorema da Multiplicação*

$$P(A_1 A_2 \dots A_N) = P(A_1)P_{A_1}(A_2)P_{A_1 A_2}(A_3) \dots P_{A_1 A_2 A_3 \dots A_{N-1}}(A_N) \quad (3.2.9)$$

Além disso temos

$$P_A(B) \geq 0 \quad (3.2.10)$$

$$P_A(E) = 1 \quad (3.2.11)$$

$$P_A(B + C) = P_A(B) + P_A(C) \quad (3.2.12)$$

Comparando as fórmulas 3.2.10, 3.2.11, 3.2.12 com os axiomas (III), (IV) e (V), podemos encontrar o sistema F de conjuntos e juntamente com o conjunto $(A | B)$ (sendo A um conjunto fixo), formam o campo de probabilidade, e portanto essas fórmulas são verdadeiras também para a probabilidade condicional $P_A(B)$. É observável também que:

$$P_A(A) = 1 \quad (3.2.13)$$

Usando 3.2.8 e outra equação análoga

$$P(A | B) = P(B)P_B(A)$$

Podemos obter uma fórmula que [7] dá certa importância em seu trabalho, pois é a essência do Teorema de Bayes:

$$P_B(A) = \frac{P(A)P_A(B)}{P(B)} \quad (3.2.14)$$

Logo a seguir falaremos mais sobre esse teorema, no momento vamos apenas analisar a mesma situação apresentada por [7] sobre a ótica da Teoria dos Conjuntos.

Quando A e B são eventos mutuamente excludentes,

$$A \cap B = \emptyset \rightarrow P(A \cup B) = P(A) + P(B)$$

quando não são mutuamente excludentes ocorre que,

$$A \cap B \neq \emptyset \rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Este conceito é consequência do chamado Princípio da inclusão e exclusão. É aí que entram os estudos de Probabilidade Condicional: ela calcula a probabilidade da interseção de dois eventos a partir de um que já ocorreu antes. Em outras palavras, o problema que ela soluciona é qual a probabilidade de um evento A ocorrer, dado que um evento B já ocorreu. O cálculo é feito através da razão entre as chances da interseção da ocorrência dos eventos A e B por as chances do evento que ocorreu primeiro, dado que no nosso exemplo um evento B já ocorreu, seria

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

Reescrevendo a equação de outra forma temos:

$$P(A \cap B) = P(A | B) \cdot P(B).$$

Podemos observar o Diagrama de Venn extraído de [6]:

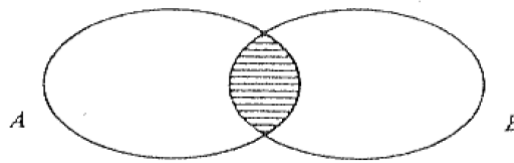


Figura 3.2.1: Diagrama de Venn $A \cap B$ retirado de [6] pg. 14

Probabilidade Condicional possui uma interpretação intuitiva em termos de frequências relativas. Pensando em probabilidade como frequência relativa 3.2.2, temos:

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{\frac{1}{n} \chi(\text{Número de ocorrências de "A e B" em } n \text{ ensaios independentes do experimento})}{\lim_{n \rightarrow \infty} \frac{1}{n} \chi(\text{Número de ocorrências de B em } n \text{ ensaios independentes do experimento})} \\ &= \frac{\lim_{n \rightarrow \infty} \frac{1}{n} \chi(\text{Número de ocorrências } A \cap B \text{ em } n \text{ ensaios})}{\lim_{n \rightarrow \infty} \frac{1}{n} \chi(\text{Número de ocorrências de B em } n \text{ ensaios})} \end{aligned} \quad (3.2.15)$$

Decorre da definição que $P(A \cap B) = P(B)P(A | B)$. Esta igualdade se generaliza quando tomamos A,B,C como eventos aleatórios, temos $P(A \cap B \cap C) = P(A)P(B | A)P(C | A \cap B)$, como podemos observar em outro diagrama:

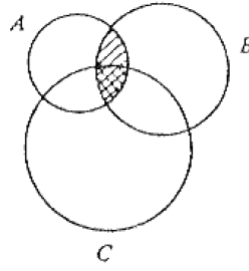


Figura 3.2.2: Diagrama de Venn $A \cap B$ retirado de [6] pg. 15

Pela prova formal temos que $P(A \cap B \cap C) = P(A \cap B)P(C | A \cap B) = P(A)P(B | A)P(C | A \cap B)$ e por indução temos o seguinte teorema

Teorema 3.1. (Teorema da Probabilidade Composta) *Seja (Ω, \mathcal{F}, P) um espaço de probabilidade. Então*

1. $P(A \cap B) = P(A)P(B | A) = P(B)P(A | B), \forall A, B \in \mathcal{F}$
2. $P(A_1 \cap A_2 \cap A_3 \dots A_n) = P(A_1)P(A_2 | A_1)P(A_3 | A_1 \cap A_2) \dots P(A_n | A_1 \cap A_2 \cap A_3 \dots A_{n-1})$
 $\forall A_1, \dots, A_n \in \mathcal{F}, \forall n = 2, 3, \dots$

O Teorema de Bayes relaciona as probabilidades de A e B com as respectivas probabilidades condicionadas mútuas. Este teorema afirma que:

$$P(A | B) = P(A | \bar{B}) \cdot \frac{P(B)}{P(A)}$$

Antes de analisar o Teorema de Bayes, vamos analisar

Teorema 3.2. *Seja $A_1 + A_2 + \dots + A_n = E$, assumindo que os eventos A_i são mutuamente excludentes e seja um X arbitrário. Então*

$$P(X) = P(A_1)P_{A_1}(X) + P(A_2)P_{A_2}(X) + \dots + P(A_n)P_{A_n}(X)$$

Demonstração.

$$X = A_1X + A_2X + \dots + A_nX;$$

Usando a equação 3.2.6

$$P(X) = P(A_1X) + P(A_2X) + \dots + P(A_nX)$$

e também utilizando a equação 3.2.8, temos

$$P(A_iX) = P(A_i)P_{A_i}(X)$$

□

Teorema 3.3. *Teorema de Bayes.* Se $A_1 + A_2 + \dots + A_n = E$ e X arbitrário, então

$$P_X(A_i) = \frac{P(A_i)P_{A_i}(X)}{P(A_1)P_{A_1}(X) + P(A_2)P_{A_2}(X) + \dots + P(A_n)P_{A_n}(X)} \quad (3.2.16)$$

$i=1,2,\dots,n$

A_1, A_2, \dots, A_n são chamadas de hipóteses e a equação 3.2.16 é considerada a probabilidade de $P_X(A_i)$ da hipótese de A_i depois da ocorrência do evento X . [$P(A_i)$ é a probabilidade de A_i]

Demonstração. Na equação que foi destacada por [7] a equação 3.2.14, temos que

$$P_X(A_i) = \frac{P(A_i)P_{A_i}(X)}{P(X)}$$

agora basta substituir a probabilidade de $P(X)$ pelo seu valor encontrado na equação 3.2. □

Exemplo 1: Numa caixa há 12 bolas, sendo 8 brancas e 4 pretas. Dois são tirados sucessivamente e sem reposição. Qual é a probabilidade de serem retiradas 2 bolas brancas? Para compreendermos melhor o cálculo vamos construir um diagrama de probabilidades mostrado na Figura 3.2.3.

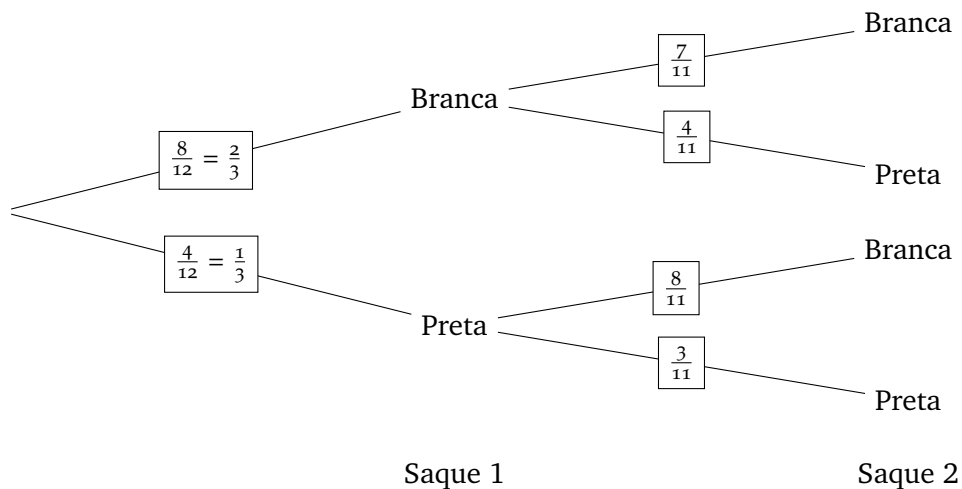


Figura 3.2.3: Diagrama de árvore com as possibilidades de escolha de livros.

A pergunta feita é $P(\text{Sacar branca} \cap \text{Sacar branca})$, portanto a resposta desejada é $P(\text{Sacar Branca no primeiro}) \cdot P(\text{Sacar branca, sabendo que já saiu branca})$, em números:

$$\underbrace{\frac{14}{33}}_{P(B \cap B)} = \underbrace{\frac{2}{3}}_{P(B)} \cdot \underbrace{\frac{7}{11}}_{P(B|B)}$$

Isso mostra que, nesse caso, mesmo que tenhamos mais bolas Brancas do que pretas, as chances de saírem duas bolas brancas seguidas é menor que 50%.

Exemplo 2: Uma ação, às terças-feiras, tem a probabilidade de 70% de subir 4% e 30% de cair 1%. Já às sextas-feiras, tem a probabilidade de 35% de subir 2% e 65% de cair 3%. Quais são as possibilidades de alta e de queda de suas respectivas probabilidades para o período de quinta à sexta-feira?

Para auxiliar na resolução do problema vamos, novamente, utilizar o diagrama de árvore mostrado na Figura 2.

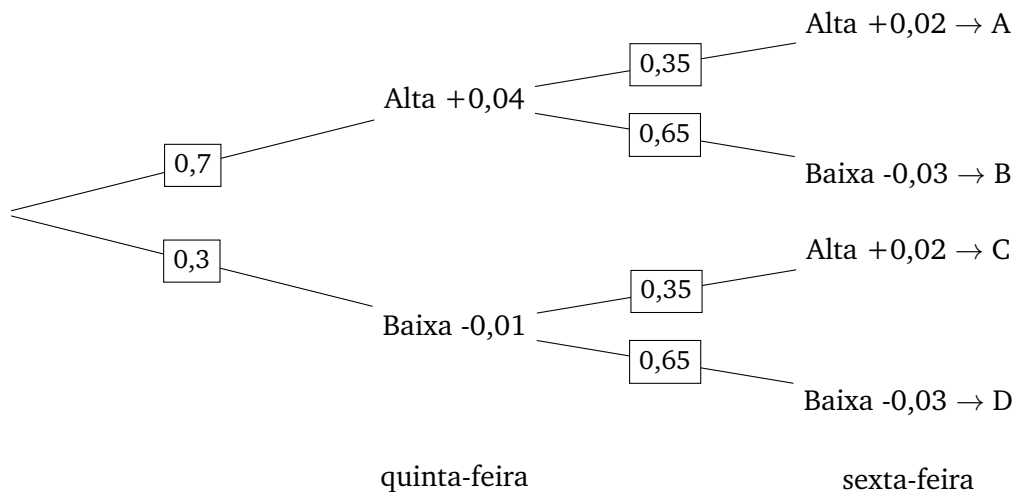


Figura 3.2.4: Diagrama de árvore com as possibilidades de alta e baixa de um ativo.

Como podemos observar, no fim da sexta-feira temos 4 possibilidades de acontecimentos, são eles:

- A)** Alta de 4% na quinta e alta de 2% na sexta. O cálculo da probabilidade é $P(+4\% \cap +2\%) = 0,7 \cdot 0,35 = 0,245$. A chance de subir na quinta e cair na sexta é de 24,5%. Como $1,04 \cdot 1,02 = 1,0608$, então tem-se 24,5% de chance da ação estar a 106,08% do preço original no fim da sexta-feira.
- B)** Alta de 4% na quinta e baixa de 3% na sexta. O cálculo da probabilidade é $P(+4\% \cap -3\%) = 0,7 \cdot 0,65 = 0,455$. A chance de subir na quinta e cair na sexta

é de 45,5%. Como $1,04 \cdot 0,97 = 1,0088$, então tem-se 45,5% de chance da ação estar a 100,88% do preço original no fim da sexta-feira.

- C) Baixa de 1% na quinta e alta de 2% na sexta. O cálculo da probabilidade é $P(-1\% \cap +2\%) = 0,3 \cdot 0,35 = 0,105$. A chance de cair na quinta e subir na sexta é de 10,5%. Como $0,99 \cdot 1,02 = 1,0098$, então tem-se 10,5% de chance da ação estar a 100,98% do preço original no fim da sexta-feira.
- D) Baixa de 1% na quinta e baixa de 3% na sexta. O cálculo da probabilidade é $P(-1\% \cap -3\%) = 0,3 \cdot 0,65 = 0,195$. A chance de cair na quinta e na sexta é de 19,5%. Como $0,99 \cdot 0,97 = 0,9603$, então tem-se 19,5% de chance da ação estar a 96,03% do preço original no fim da sexta-feira.

3.2.4 Independência

Definição 4. Seja (Ω, \mathcal{F}, P) um espaço de probabilidade. Os eventos aleatórios A e B são (estocasticamente) independentes se:

$$P(A \cap B) = P(A)P(B) \quad (3.2.17)$$

Proposição 2. A é independente de si mesmo se, e somente se, $P(A) = 0$ ou $P(A) = 1$

Demonstração. $P(A) = P(A \cap A) = P(A)P(A) \leftrightarrow P(A) = 0$ ou $P(A) = 1$ □

Proposição 3. Se A e B são independentes, então A e B^c também são independentes (e também A^c e B , e ainda A^c e B^c).

Demonstração. Supondo A, B independentes. Então $P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(B^c)$. □

Observação 3.2.1. Quando $P(A \cap B) = 0$, não dizemos que A e B são independentes, mas que são incompatíveis. Veja que se num lançamento de uma moeda $P(A) = \text{“sair cara”}$ e $P(B) = \text{“sair coroa”}$, então $P(A \cap A^c) = 0 \neq \frac{1}{4} = P(A)P(A^c)$

Definição 5. Os eventos aleatórios $A_i, i \in I$, são independentes 2 a 2 se

$$P(A_i \cap A_j) = P(A_i)P(A_j) \forall i, j \in I, i \neq j \quad (3.2.18)$$

Definição 6. a) Os eventos A_1, A_2, \dots, A_n são chamados estocasticamente independentes se:

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_m}) = P(A_{i_1})P(A_{i_2})\dots P(A_{i_m}) \quad (3.2.19)$$

$$\forall 1 \leq i_1 \leq i_2 \leq \dots \leq i_m \leq n, \forall 2 \leq m \leq n$$

b) Os eventos A_1, A_2, \dots são independentes se $\forall n \geq 2$, A_1, A_2, \dots, A_n são independentes.

c) Os eventos $A_i, i \in I$ [onde I é um conjunto de índices tal que $I \geq 2$] são independentes se toda a subfamília finita deles é de eventos independentes, i.e. se $A_{i_1}, A_{i_2}, \dots, A_{i_m}$ são independentes para toda combinação i_1, i_2, \dots, i_m de elementos de $I, \forall m=2,3,\dots$

Proposição 4. Se os eventos $A_i, i \in I$, são independentes, então os eventos $B_i, i \in I$, também são independentes, onde cada B_i , é igual a um A_i , ou A_i^c .

Demonstração. Pelos itens (a) e (c) da definição 6, basta provar que toda subfamília finita dos B_i , satisfaz a regra do produto. \square

3.2.5 Covariância

Para estudar a dependência entre duas variáveis, podemos calcular sua covariância, que é uma média de como duas variáveis variam conjuntamente. Seus valores se situam no intervalo $(-\infty, +\infty)$; se for positiva significa que existe uma relação diretamente proporcional entre elas e se for negativa implica que elas são inversamente proporcionais. A covariância entre duas variáveis é medida através da seguinte equação:

$$\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y] \quad (3.2.20)$$

Quando $\text{cov}(X, Y) = 0$, significa que essas variáveis não são correlacionadas.

3.2.6 Função Densidade de Probabilidade

Não confundir Função Probabilidade com Função Densidade de Probabilidade, visto que a Função Probabilidade trabalha com valores DISCRETOS, enquanto a F.D.P. trata

de valores CONTÍNUOS. Além disso é uma função não negativa, a integral da função no espaço amostral é igual a um e a probabilidade de ocorrência de um evento é a sua integral entre os limites da variável aleatória que define esse evento. Nesse caso imaginemos que um determinado dia esperamos certa quantidade de chuva de acordo com o gráfico abaixo dado pelo sistema:

$$f(x) = \begin{cases} 0, & \text{se } x < 0 \\ 0.04x, & \text{se } 0 \leq x \leq 5 \\ -0.04 + 0.4, & \text{se } 5 < x \leq 10 \\ 0, & \text{se } 10 < x \end{cases}$$

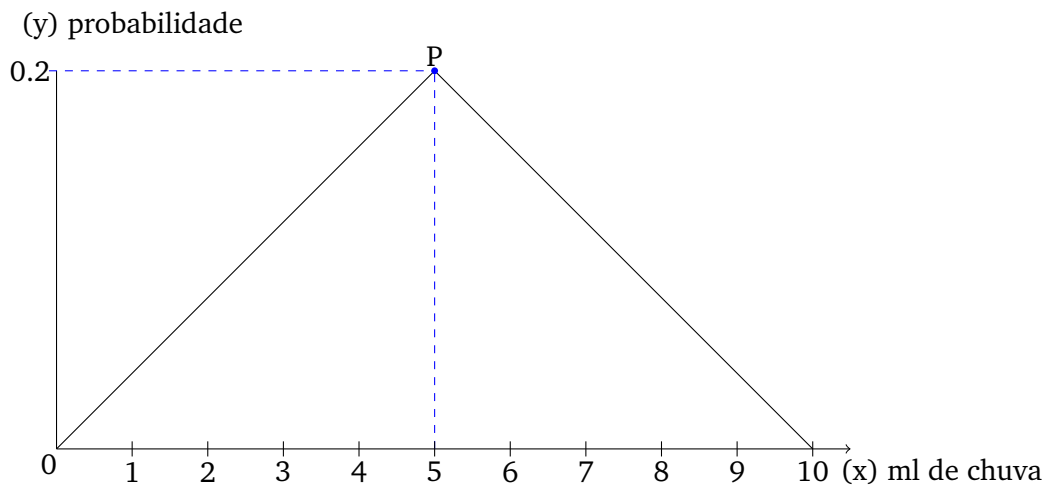


Figura 3.2.5: Gráfico da F.D.P. da quantidade de chuva

O fato é que se perguntarmos qual a probabilidade de chover exatamente 5 ml, de forma errônea, porém intuitiva, dizemos que é 0.2 ou 20% quando na verdade a chance de chover exatamente 5 ml é zero, pois a variável é contínua e portanto sempre choverá um pouco a mais ou a menos. Para podermos calcular a chance de algum acontecimento, precisamos escolher um intervalo de continuidade. O correto então é perguntar quais são as chances de chover entre [4 , 6] ml.

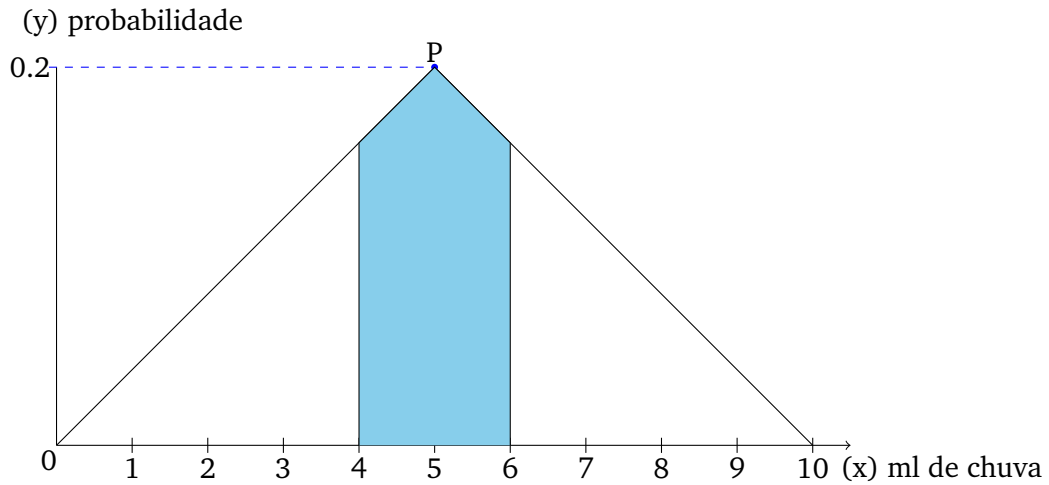


Figura 3.2.6: Gráfico da F.D.P. da quantidade de chuva de 4 a 6 ml

Ou ainda poderíamos aproximar ainda mais de 5 ml escolhendo os valores 4.99 e 5.01, contudo quanto mais aproximamos de um valor exato, menor será a área criada no gráfico e portanto vai diminuindo, veja nas figuras 3.2.7, 3.2.8 e 3.2.9.

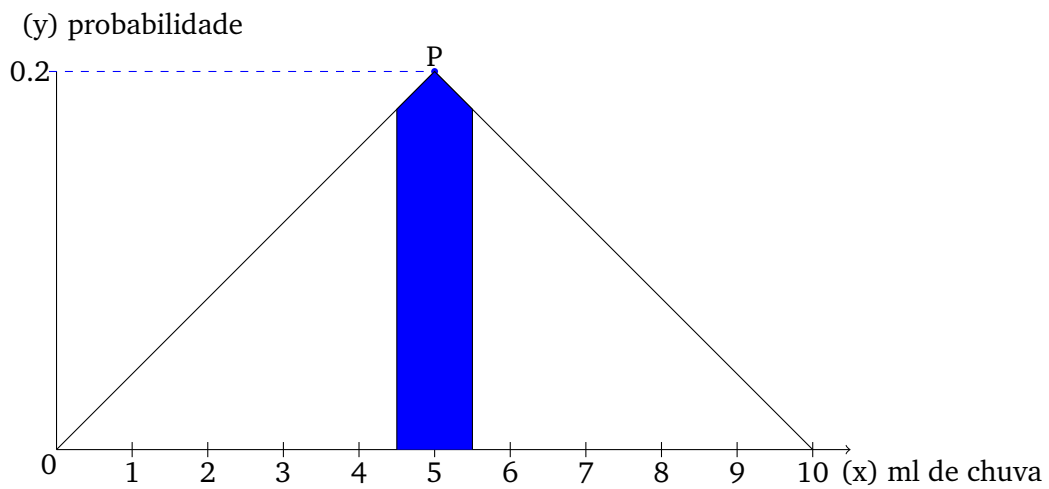


Figura 3.2.7: Gráfico da F.D.P. da quantidade de chuva de 4.5 a 5.5 ml

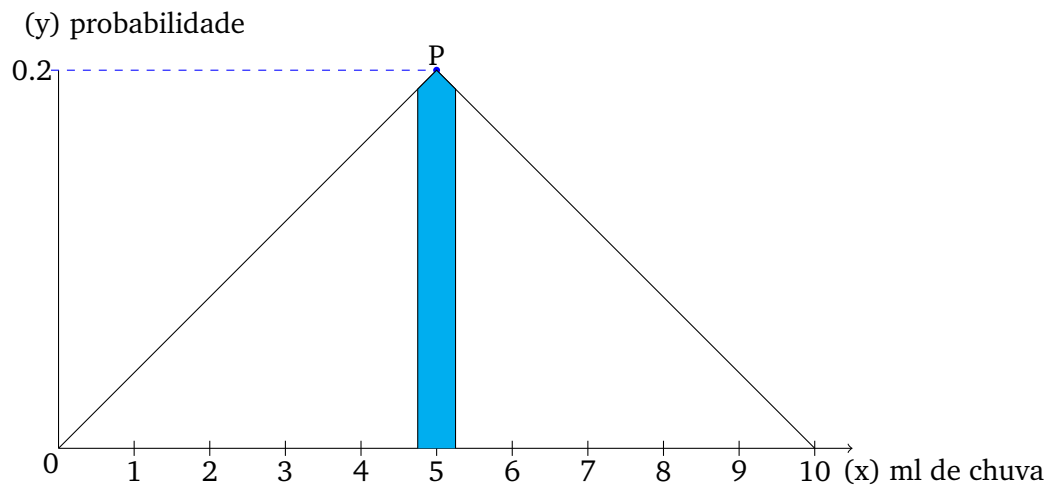


Figura 3.2.8: Gráfico da F.D.P. da quantidade de chuva de 4,75 a 5,25

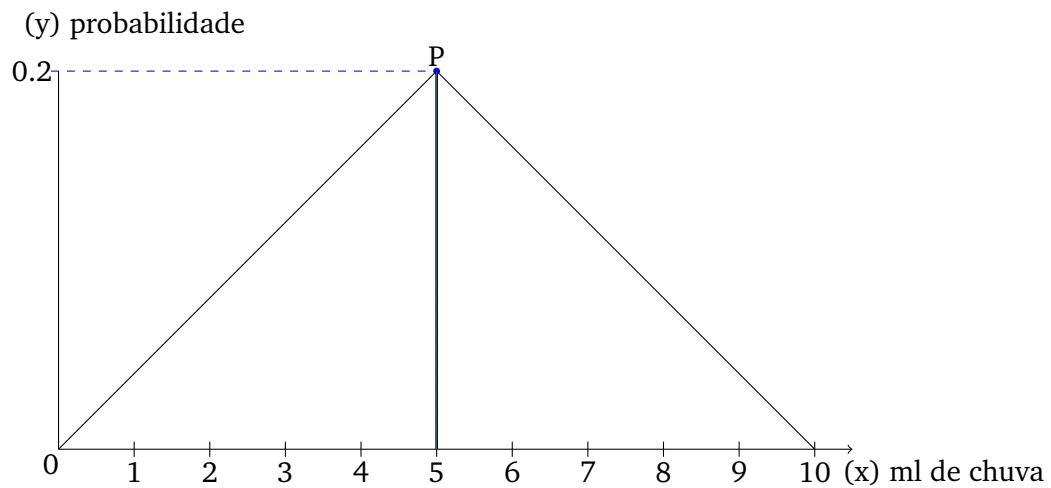


Figura 3.2.9: Gráfico da F.D.P. da quantidade de chuva de 4,99 a 5,01

Como podemos observar na figura 3.2.9 a área da figura formada é muito pequena e conforme vamos nos aproximando do valor exato a chance de tal acontecimento tende a zero. Podemos perguntar também $P(0 < x \leq 10)$ e o gráfico da figura 3.2.10 mostra o resultado em azul.

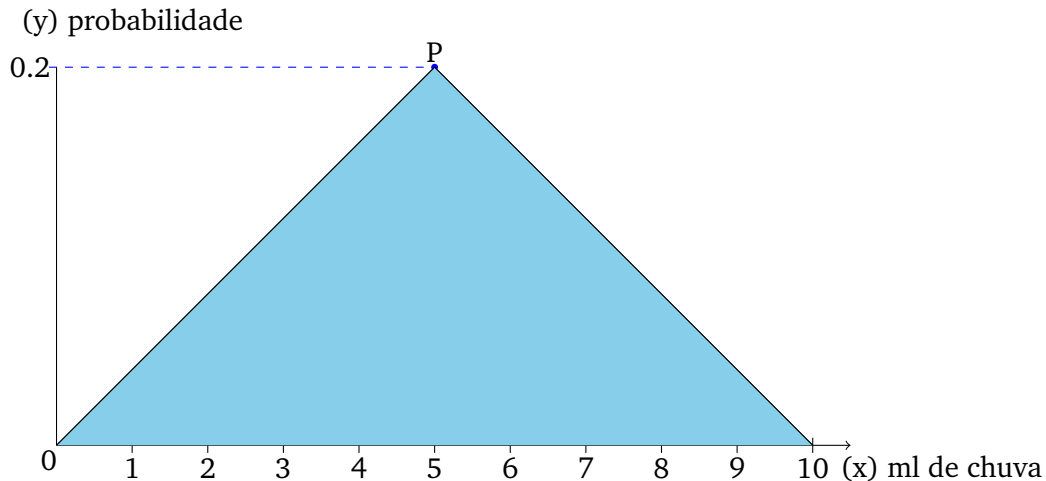


Figura 3.2.10: Gráfico da F.D.P. da quantidade de chuva de 0 a 10

Devemos perceber que essa é a área total onde se incluem todos os possíveis acontecimentos e portanto podemos defini-la também como nosso *Espaço Amostral*. Agora vamos calcular as chances de chover até 1 ml ou seja $P(0 < x < 1)$. Para calcular a área de gráficos precisamos realizar o processo de integração, vamos a ele:

$$\begin{aligned} P(0 < x < 1) &= \int_0^1 f(x) dx = \\ &= \int_0^1 0.04x dx = \\ &= \left|_0^1 0.02x^2 \right. = \\ &= 0.02, \end{aligned}$$

ou seja, as chances de chover até 1 ml são de 2%. Vejamos agora as chances de $P(4 < x < 6)$, para resolver esse problema vamos enxergar que o gráfico é simétrico no valor 5 ml e portanto podemos calcular apenas, $\frac{1}{2} - P(x < 4)$ e multiplicar por 2.

$$\begin{aligned} P(4 < x < 6) &= \frac{1}{2} - \int_0^4 f(x) dx = \\ &= \frac{1}{2} - \int_0^4 0.04x dx = \\ &= \frac{1}{2} - \left|_0^4 0.02x^2 \right. = \\ &= 0.5 - 0.32 = \\ &= 0.18. \end{aligned}$$

Portanto chover entre 4 e 6 ml acontecerá em 36% das vezes ou temos 0.36 de segurança que choverá entre 4 e 6. Para calcular as chances de um acontecimento não discreto, devemos usar então:

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x) dx \quad (3.2.21)$$

Obedecendo sempre, é claro, as regras de integração. Vale perceber que $F(x)$ é a função utilizada, P é a probabilidade de ocorra os eventos no intervalo $[x_1, x_2]$. A função probabilidade não resolve problemas de função densidade, mas o inverso ocorre. Utilizando o mesmo raciocínio, para valores discretos podemos escrever a equação 3.2.21 do seguinte modo:

$$P(x_1, x_2, \dots, x_n) = F(x_1) + F(x_2) + \dots + F(x_n) = \sum_{i=1}^n F(x_i)$$

Estendendo o pensamento, ainda em valores discretos, podemos perceber que a variação da probabilidade é o produto de uma função em x conforme x varia:

$$\Delta P = F(x) \cdot \Delta x \quad (3.2.22)$$

3.3 ESTATÍSTICA

Estatística é a ciência que se utiliza das teorias probabilísticas para explicar a frequência da ocorrência de eventos, tanto em estudos observacionais quanto em experimentos para modelar a aleatoriedade e a incerteza de forma a estimar ou possibilitar a previsão de fenômenos futuros, conforme o caso.

Algumas práticas estatísticas incluem, por exemplo, o planejamento, a sumarização e a interpretação de observações. Dado que o objetivo da estatística é a produção da melhor informação possível a partir dos dados disponíveis, alguns autores sugerem que a estatística é um ramo da teoria da decisão.

A estatística é uma ciência que se dedica à coleta, análise e interpretação de dados. Preocupa-se com os métodos de recolha, organização, resumo, apresentação e interpretação dos dados, assim como tirar conclusões sobre as características das fontes donde estes foram retirados, para melhor compreender as situações.

Esta é a ciência que nos ajudará a determinar a Volatilidade do preço de uma ação.

A Estatística trabalha com dois tipos de variáveis. As variáveis qualitativas, quando não são associadas a número: Estado civil, profissão, escolaridade entre outros. Já as variáveis quantitativas são associadas a números, como: idade, altura, peso e também preço, que o tema deste trabalho.

3.3.1 Tabela de Frequência

Para as variáveis quantitativas, constroem-se tabelas de frequência. Observe o exemplo a seguir:

Tabela 3.3.1: Exemplo de Frequência absoluta e relativa

Profissão	Frequência Absoluta	Frequência Relativa	Porcentagem %
Professor	30	$\frac{30}{50}=0,6$	60
Pesquisador	15	$\frac{15}{50}=0,3$	30
Investidor	5	$\frac{5}{50}=0,1$	10
Total	50	1	100

As variáveis trabalhadas na tabela 3.3.1 são discretas. A partir da tabela, pode-se afirmar que a probabilidade de um indivíduo ser professor é de $P(\text{professor})=60\%$

Vamos agora construir uma tabela de frequência para variáveis contínuas que será a tabela 3.3.2, como por exemplo a renda de algumas famílias.

Tabela 3.3.2: Exemplo de Classes Frequência absoluta e relativa

Renda em R\$	Frequência Absoluta	Frequência Relativa	Porcentagem %
400 † 700	90	$\frac{90}{200}=0,45$	45
700 † 1500	80	$\frac{70}{200}=0,4$	40
1500 † 2200	30	$\frac{30}{200}=0,15$	15
Total	200	1	100

Como vimos, as tabelas para variáveis discretas e contínuas são diferentes. Nas variáveis contínuas usamos o símbolo † para representar classes de valores, pois ficaria inviável dizer a frequência absoluta ou relativa de cada valor dentro de uma classe. Assim de acordo com a última tabela, a porcentagem de famílias que ganham até R\$1500,00 é de 85%.

3.3.2 Medidas de Centralidade

Além de determinar a probabilidade de um evento ocorrer através do levantamento frequencial, podemos também determinar qual é o "valor mais esperado" ou o "valor verdadeiro" que represente um conjunto de medidas. Esse valor é chamado de "valor central" ou "valor médio" e pode se determiná-lo de várias maneiras, vejamos a seguir.

Mediana

Sejam $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_n$ os n valores de uma variável X . Define-se como mediana - denotado por Me :

- se n é ímpar, então a mediana será o elemento $x_{(\frac{n+1}{2})}$;
- se n é par, então a mediana será o valor da expressão $\frac{x_{(\frac{n}{2})} + x_{(1+\frac{n}{2})}}{2}$.

Vejamos o exemplo a seguir:

- Um estudo para verificar a eficiência, em termos de horas de duração, de algumas lâmpadas coletou o seguinte resultado:

5 - 4 - 9 - 6 - 3 - 8 - 1 - 8 - 6 - 8 - 3 - 11

Para encontrar a Me precisamos ordenar a lista de valores. portanto:

1 - 3 - 3 - 4 - 5 - 6 - 6 - 8 - 8 - 8 - 9 - 11

Agora que $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_{12}$, podemos calcular a Me . Como temos um número par de elementos vamos fazer o seguinte:

$$\frac{x_6 + x_7}{2} = \frac{6+6}{2} = 6$$

Assim a mediana do funcionamento dessas lâmpadas é de 6 horas.

Moda

É a ocorrência que tem a maior frequência entre os valores de um conjunto. Vejamos a moda do exemplo anterior: denota-se Mo .

5 - 4 - 9 - 6 - 3 - 8 - 1 - 8 - 6 - 8 - 3 - 11

Claro que podemos ordená-los, mas não é necessário. Agora analisando, vemos que o valor mais frequente é 8. Portanto, por M_o , o tempo de duração de uma lâmpada é de 8 horas.

Média Aritmética

Sejam $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_n$ os valores de n elementos de um conjunto. Define-se a média aritmética - denotada por \bar{x} - como a razão entre a soma de seus elementos e o número total de elementos:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

Somente para deixar mais implícito,

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i \quad (3.3.1)$$

Apliquemos agora no exemplo anterior.

$$\bar{x} = \frac{5+4+9+6+3+8+1+8+6+8+3+11}{12} = \frac{72}{12} = 6$$

Assim, por média aritmética temos que o tempo de funcionamento de uma lâmpada dessa amostra é de 6 horas.

Agora que entendemos o funcionamento da média aritmética voltemos às funções densidade probabilidade. Relembrando a equação 3.2.22, podemos extrair a média aritmética:

$$\Delta P = F(x)\Delta x$$

Equação 3.2.22, que calculava a probabilidade a partir do produto de uma função $F(x)$ com a variação do x (Δx)

Equação 3.3.1, quando x_i não varia, sendo x .

$$\bar{x} = \frac{1}{n} \cdot \sum x \cdot n$$

(Ex: probabilidade de sair cara em uma moeda não viciada.) $\bar{x} = \frac{1}{n} \cdot \sum x \cdot n \cdot F(x)\Delta x$

A última equação é a média aritmética da probabilidade da soma sobre todos os intervalos Δx possíveis e é válida para qualquer FDP. Apenas simplificando obtemos:

$$\bar{x} = \sum x \cdot F(x)\Delta x \quad (3.3.2)$$

Média Geométrica

Para muitos casos a média aritmética é extremamente útil, porém para outros como taxa de juros por um período de tempo, ela perde efeito e acaba não apontando o resultado mais coerente. Para este tipo de problema temos a média geométrica. Vejamos a tabela 3.3.3

Tabela 3.3.3: Histórico de taxas

Período (mês)	Taxa (Juros)	Índice de Correção
1	100%	2
2	2%	1,02
3	2%	1,02
4	1%	1,01
5	10%	1,1

Repare que pela fórmula de juros compostos 3.1.5, o Montante acumulado será:

$$C_5 = C_0 \cdot 2 \cdot 1,02 \cdot 1,02 \cdot 1,01 \cdot 1,1 = C_0 \cdot 2,3117$$

Assim podemos afirmar que o rendimento, passado cinco meses, foi de 131,17%. Para calcular a taxa média mensal não podemos utilizar a média aritmética, veja que $\bar{x} = 23\%$, o que é incorreto, já que

$$C_0 \cdot (1,23)^5 = C_0 \cdot 2,8153,$$

que difere do valor real. Portanto vamos utilizar a média geométrica - denotada por x_g - que ao invés de dividir pelo total de n termos, faz a raiz enésima da multiplicação de todos os termos:

$$x_g = \sqrt[n]{x_1 \cdot x_2 \cdots x_n} = \sqrt[n]{\prod_{i=1}^n x_i}$$

Agora vamos aplicar a x_g :

$$\sqrt[5]{2 \cdot 1,02 \cdot 1,02 \cdot 1,01 \cdot 1,1} = \sqrt[5]{2,3117} = 1,1824$$

Portanto a taxa de juros é de 18,24% e aplicando isso a taxa de juros compostos 3.1.5, temos:

$$C_o \cdot (1,1824)^5 = C_o \cdot 2,3117$$

obtendo assim uma resposta coerente.

3.3.3 Medidas de Dispersão

As medidas de centralidade nos ajudam a encontrar um valor mais esperado de uma situação; contudo, mesmo o chamando de valor verdadeiro, não é de fato sempre que ele ocorre. Voltemos ao exemplo das lâmpadas que tinha $Me=6$, $Mo=8$, $\bar{x}=6$ e façamos a seguinte pergunta: qual a probabilidade de retirarmos uma lâmpada que dure 6 ou 8 horas? Antes de responder esta questão, vamos analisar que estamos considerando os valores verdadeiros de dois modos diferentes portando a resposta deveria ser 100%. Agora vamos aos cálculos:

Todas as possibilidades = 1 - 3 - 3 - 4 - 5 - 6 - 6 - 8 - 8 - 8 - 9 - 11

Possibilidades preferíveis = 6 - 6 - 8 - 8 - 8, por 3.2.1

$$\frac{5}{12} = 0,4166$$

Ou seja, as chances de escolher uma lâmpada que se encontrava dentro dos valores mais prováveis são menores que 50%. Isso se deve a um fator chamado de volatilidade. Em resumo, a volatilidade é o quanto as medidas de um conjunto tendem a se afastar das medidas centrais e no nosso caso das lâmpadas podemos afirmar que a volatilidade é bem alta ou seja, temos uma dispersão muito grande dos valores centrais, os valores tem uma variação elevada em relação às medidas centrais e nesses casos as medidas centrais perdem utilidade. Vamos agora a algumas definições e métodos de cálculo de dispersão, para assim podermos dar mais precisão ao valor esperado.

Desvio

Considere um conjunto de valores cujo valor esperado seja x_0 . Dada uma medida x_1 , define-se Desvio da medida central como:

$$d = x_1 - x_0.$$

Desvio Médio

Considerando um conjunto de valores, pode-se calcular a média dos desvios fazendo a diferença entre os valores agrupados dois a dois, do primeiro ao último:

$$d_{i,j} = x_i - x_j.$$

Observe a tabela 3.3.4:

Tabela 3.3.4: Tabela de desvio entre todos os valores

$x_i - x_j$	x_1	x_2	x_3	\dots	x_n
x_1	0	$d_{1,2}$	$d_{1,3}$	\dots	$d_{1,n}$
x_2	$d_{2,1}$	0	$d_{2,3}$	\dots	$d_{2,n}$
x_3	$d_{3,1}$	$d_{3,2}$	0	\dots	$d_{3,n}$
\vdots	\vdots	\vdots	\vdots	0	\vdots
x_n	$d_{n,1}$	$d_{n,2}$	$d_{n,3}$	\dots	0

Vamos contar o número de elementos pertencentes a esse conjunto:

$n \cdot n$, (número de linhas vezes o número de colunas.)

$n \cdot n - n$, (menos as diagonais que são 0.)

$$n(n - 1)$$

$$\frac{n(n-1)}{2}, \text{ pois os módulos } |d_{i,j}| = |d_{j,i}|$$

Então, um modo de medir o quão espalhados os valores estão de um *Valor esperado* é:

$$\bar{d} = \frac{1}{2 \cdot n \cdot (n - 1)} \cdot \sum_{i,j} d_{i,j}. \quad (3.3.3)$$

Porém, observe que i e j variam independentemente de 1 a n ; tem-se que $d_{i,j} = -d_{j,i}$, tornando o somatório sempre nulo.

Variância

Um modo de evitar o anulamento é usar o quadrado da diferença; veja na tabela 3.3.5:

Tabela 3.3.5: Tabela de diferenças ao quadrado

$x_i - x_j$	x_1	x_2	x_3	\dots	x_n
x_1	0	$d_{1,2}^2$	$d_{1,3}^2$	\dots	$d_{1,n}^2$
x_2	$d_{2,1}^2$	0	$d_{2,3}^2$	\dots	$d_{2,n}^2$
x_3	$d_{3,1}^2$	$d_{3,2}^2$	0	\dots	$d_{3,n}^2$
\vdots	\vdots	\vdots	\vdots	0 \dots 0	\vdots
x_n	$d_{n,1}^2$	$d_{n,2}^2$	$d_{n,3}^2$	\dots	0

Com o mesmo raciocínio da soma do Desvio médio agora temos

$$\bar{d}^2 = \frac{1}{2 \cdot n \cdot (n - 1)} \cdot \sum_{i,j} d_{i,j}^2. \quad (3.3.4)$$

Assim obtemos a medida do quadrado da dispersão dos dados. E ainda vale destacar que

$$d_{i,j} = x_i - x - j = x_i - x_j + 0 = x_i - x_j + \bar{x} - \bar{x} = (x_i - \bar{x}) - (x_j - \bar{x}) = d_i - d_j,$$

onde fica definido que

$$d_i = x_i - \bar{x}.$$

Substituindo na equação 3.3.4:

$$\begin{aligned}\bar{d}^2 &= \frac{1}{2 \cdot n \cdot (n-1)} \cdot \sum_{i,j} (d_i - d_j)^2 = \\ &= \frac{1}{2 \cdot n \cdot (n-1)} \cdot \sum_{i,j} (d_i^2 + d_j^2 - 2 \cdot d_i \cdot d_j) = \\ &= \frac{1}{2 \cdot n \cdot (n-1)} \left[\sum_{i,j} d_i^2 + \sum_{i,j} d_j^2 - 2 \cdot \sum_{i,j} d_i \cdot \sum_{i,j} d_j \right]\end{aligned}$$

Como

$$\sum_i d_i = 0 \text{ e } \sum_j d_j = 0 \text{ segue que, } 2 \cdot \sum_{i,j} d_i \cdot \sum_{i,j} d_j = 0 \text{ e portanto}$$

$$\bar{d}^2 = \frac{1}{2 \cdot n \cdot (n-1)} \left[\sum_{i,j} d_i^2 + \sum_{i,j} d_j^2 \right].$$

Como a soma em i é igual em j ,

$$\sum_{i,j} d_i^2 = \sum_{i,j} d_j^2.$$

Temos que

$$\begin{aligned}\sum_{i,j} d_i^2 + \sum_{i,j} d_j^2 &= 2 \cdot \sum_{i,j} d_i^2 \\ 2 \cdot \sum_{i,j} d_i^2 &= 2 \cdot \sum_i \sum_j d_i^2 \\ &= 2 \cdot n \cdot \sum_i d_i^2.\end{aligned}$$

Voltando a equação 3.3.4 do desvio ao quadrado obtemos

$$\begin{aligned}\bar{d}^2 &= \frac{1}{2 \cdot n \cdot (n-1)} \cdot \sum_{i,j} d_{i,j}^2 \\ &= \frac{1}{2 \cdot n \cdot (n-1)} \cdot 2 \cdot n \cdot \sum_i d_i^2,\end{aligned}$$

resultando no que chamamos de Variância Amostral:

$$S^2 = \frac{1}{n-1} \sum_i d_i^2 \quad (3.3.5)$$

Destacamos também a equação da Variância Populacional:

$$\sigma^2 = \frac{1}{N} \sum_i (x_i - \bar{x})^2 \quad (3.3.6)$$

Sendo N a população, a diferença entre as Variâncias no denominador - n e $(n-1)$ - ocorre, pois quando calculamos a Variância amostral usamos a média das amostras que não é necessariamente a média da população. O que ocorre é que se dividirmos por n , estamos subestimando o desvio amostral em relação ao desvio populacional, quando dividimos por $(n-1)$ nós fazemos com que o valor da variância amostral aproxime da variância populacional.

Desvio Padrão

Para redimensionar o valor da variância, calcula-se a raiz quadrada e assim obtemos o desvio padrão:

$$S = \sqrt{S^2} = \sqrt{\frac{1}{n-1} \cdot \sum_i (x_i - \bar{x})^2} \quad (3.3.7)$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{n} \cdot \sum_i (x_i - \bar{x})^2} \quad (3.3.8)$$

3.3.4 Distribuição Binomial

Supondo um experimento aleatório com somente dois resultados possíveis. Exemplos:

- Cara ou coroa de uma moeda;
- Acertar uma bola na cesta [sim ou não];
- Pertinência de um elemento a um conjunto $\Omega = \{\in, \notin\}$.

Todos esses resultados são excludentes. Se p é a chance de ocorrer o evento A e q a chance de ocorrer o evento B , pode se afirmar que

$$p + q = 1$$

Ao repetir N vezes o mesmo experimento e chamar de n o número de ocorrências do evento tipo A , é claro que o número de ocorrências do evento B será $m = N-n$.

Assim n é um número inteiro entre 0 e N . Como o resultado de cada experimento é independente do experimento anterior segue que a chance de um evento A ocorrer seja p .

$$P_{1,p}(1) = p, N=1 \text{ e } n=1.$$

E que para ocorra o evento A em dois experimentos é $p \cdot p$:

$$P_{2,p}(2) = \overbrace{p \cdot p}^{2 \text{ eventos do tipo } A} = p^2, N=2 \text{ e } n=2.$$

A probabilidade para que todos os eventos sejam do tipo A em N experimentos será:

$$P_{N,p}(N) = \overbrace{p \cdot p \cdot p \cdots p}^{N \text{ eventos do tipo } A} = p^N.$$

Para N experimentos onde $N-1$ ocorrem os eventos do tipo A e o último evento sendo do tipo B , poderia ser calculado da seguinte forma:

$$p \cdot p \cdot p \cdot p \cdots p \cdot q = p^{N-1} \cdot q.$$

Porém, devemos observar que ela calcula exatamente a chance de o último evento ser do tipo B e não a chance de apenas um evento B , pois o evento B pode aparecer em qualquer um dos N eventos. Assim para calcular as chances de ocorrer apenas um evento B , em qualquer uma das posições, devemos somar as chances do evento B ocorrer cada em cada uma das posições:

$$\begin{aligned} P_{N,p}(N-1) &= p \cdot p \cdot p \cdots p \cdot q \\ &+ p \cdot p \cdot p \cdots q \cdot p \\ &+ \vdots \\ &+ p \cdot q \cdot p \cdots p \\ &+ q \cdot p \cdot p \cdots p. \end{aligned}$$

Como o valor de cada parcela é $p^{N-1} \cdot q$ e como há N parcelas, o valor da probabilidade de ocorrer exatamente um evento B em N experimentos, não importando em qual experimento ele ocorra, é

$$P_{N,p}(N-1) = N \cdot p^{N-1} \cdot q.$$

De outra forma, trata-se de uma permutação de N elementos com $N-1$ repetições para o evento A e uma única ocorrência do evento B :

$$P_N^{(N-1,1)} = \frac{N!}{(N-1)! \cdot 1!} = \frac{N \cdot (N-1)!}{(N-1)! \cdot 1!} = N \text{ parcelas.}$$

Para calcular a probabilidade de N eventos do tipo A ocorrerem em N experimentos é

$$\underbrace{p \cdot p \cdot p \cdots p}_{n \text{ eventos do tipo A}} \cdot \underbrace{q \cdot q \cdot q \cdots q}_{m \text{ eventos do tipo B}} = p^n \cdot q^m = p^n \cdot q^{N-n}.$$

Agora para calcular o número de arranjos, basta calcular o número de permutações de N elementos com n repetições do evento A e $N-n$ repetições do evento B:

$$P_N^{(n, N-n)} = \frac{N!}{n! \cdot (N-n)!} \text{ parcelas.}$$

Assim obtém-se uma expressão para a probabilidade de dois eventos mutuamente excludentes e exaustivos:

$$P_{N,p}(n) = \frac{N!}{n! \cdot (N-n)!} \cdot p^n \cdot q^{N-n}. \quad (3.3.9)$$

A expressão 3.3.9 é chamada Distribuição Binomial porque, para cada n , corresponde a um termo da expansão do binômio

$$(p + q)^N = \sum_{n=0}^N \frac{N!}{n! \cdot (N-n)!} \cdot p^n \cdot q^{N-n}.$$

Lembrando que $p+q=1$,

$$\sum_{n=0}^N P_{N,p}(n) = \sum_{n=0}^N \frac{N!}{n! \cdot (N-n)!} \cdot p^n \cdot q^{N-n} = 1. \quad (3.3.10)$$

Média da Distribuição Binomial

Abordemos pelo contexto dos esportes. Uma empresa de Estatística faz um estudo sobre um exímio jogador de futebol que faz gol em 30% das vezes que a bola chega a seus pés. Se a bola chegou 13 vezes no seu pé, quantos gols espera-se que ele faça? O cálculo intuitivo, de fato, é simples:

$$\bar{n} = 13 \cdot 0,3 = 3,9.$$

Espera-se que ele faça 3,9 gols; como gol é uma variável discreta, aceita-se como resposta 3 ou 4 gols. Porém vamos agora verificar a validade desse pensamento intuitivo. Para isso utilizaremos a equação 3.3.2:

$$\bar{x} = \sum x \cdot F(x) \Delta x$$

Nosso $F(x) = P_{N,p}(n)$, portanto:

$$\bar{n} = \sum n \cdot P_{N,p} \cdot \Delta n.$$

$\Delta n = 1$, pois o número total de vezes que a bola chega ao pé do jogador é uma variável discreta. A letra \bar{n} , significa quantos gols esperamos do jogador em N vezes que a bola chegou a seus pés. Temos,

$$\begin{aligned} \bar{n} &= \sum_{n=0}^N n \cdot \frac{N!}{n!(N-n)!} \cdot p^n \cdot q^{N-n} \\ &= \sum_{n=0}^N n \cdot \frac{N!}{n \cdot (n-1)!(N-n)!} \cdot p^n \cdot q^{N-n} \\ &= \sum_{n=1}^N \frac{N!}{(n-1)!(N-n)!} \cdot p^n \cdot q^{N-n} \\ &= \sum_{n=1}^N \frac{N \cdot (N-1)!}{(n-1)!(N-n)!} \cdot p \cdot p^{n-1} \cdot q^{N-n} \\ &= N \cdot p \cdot \sum_{n=1}^N \frac{(N-1)!}{(n-1)!(N-n)!} \cdot p^{n-1} \cdot q^{N-n}. \end{aligned}$$

Repare que, se:

$$a = N - 1 \text{ ou } N = a + 1$$

e

$$b = n - 1 \text{ ou } n = b + 1.$$

Temos:

$$\bar{n} = N \cdot p \cdot \sum_{b=0}^{a+1} \frac{a!}{(b)!(a+1-(b+1))!} \cdot p^b \cdot q^{a+1-(b+1)} = N \cdot p \cdot \sum_{b=0}^{a+1} \frac{a!}{(b)!(a-b)!} \cdot p^b \cdot q^{(a-b)}.$$

Note que na equação 3.3.10 o resultado é 1. Aqui temos o mesmo exemplo:

$$\sum_{b=0}^{a+1} \frac{a!}{(b)!(a-b)!} \cdot p^b \cdot q^{(a-b)} = 1.$$

Assim provamos que basta multiplicar N por p para saber o número de acertos esperados em N tentativas dado a probabilidade p do acerto:

$$\begin{aligned}
 &= N \cdot p \cdot \overbrace{\sum_{b=0}^{a+1} \frac{a!}{(b)! \cdot (a-b)!} \cdot p^b \cdot q^{(a-b)}}^1 \\
 &= N \cdot p \cdot 1 \\
 &= N \cdot p. \tag{3.3.11}
 \end{aligned}$$

Desvio Padrão da Distribuição Binomial

Ainda assim, uma nova pergunta poderia ser feita: se a média foi de 3.9 gols, até quantos gols ainda seria comum ele fazer? Veja que apostar que o jogador fará gol em todas as vezes que a bola chegar a seus pés é apostar fora das probabilidades e assim até quantos gols seria apostar com racionalidade? Vamos supor duas situações em que o jogador que foi contratado para fazer gol em 30% das vezes que a bola chega seus pés foi exposto. A primeira situação foi em uma partida que nosso exímio jogador fez apenas 2 gols das 13 vezes que chegou à bola, assim sendo foi muito criticado pela torcida e pelos dirigentes, pois a expectativa era em torno de 3 ou 4 gols. A segunda situação acontece com o jogador ressentido e dá tudo de si no jogo, novamente a bola chega 13 vezes aos seus pés e ele faz 5 gols. A torcida o perdoa, mas os dirigentes não ficam tão contentes assim. Podemos dizer que o jogador superou as expectativas com os 5 gols? Pelo cálculo anterior, sim. Porém temos que lembrar que há variações dentro dos 30%. Os dirigentes não ficaram tão contentes, pois além de usarem a média da distribuição normal, também usaram o desvio padrão binomial:

$$s = \sqrt{n \cdot q} \quad s = \sqrt{3.9 \cdot 0.7} \approx 1.65$$

Assim, os resultados esperados para o jogador seriam:

$$[3.9 \pm 1.65] = [2.25; 5.5]$$

Os dirigentes usaram esse intervalo como intervalo de satisfação e assim 2 gols tendo recebido a bola 13 vezes, está abaixo do padrão pelo qual o jogador foi contratado, ao mesmo tempo 5 gols está dentro da satisfação esperada. Para o jogador estar acima da

média esperada teria que ter feito, no mínimo, 6 gols. Agora verifiquemos a validade desta equação.

Para demonstrar vamos fazer o mesmo procedimento anterior, porém com a variância. $F(x)$ continua sendo $P_{N,p}(n)$ e assim obtemos:

$$s^2 = \sum_{n=0}^N (n_i - \bar{x})^2 \cdot P_{N,p}(n) \cdot \Delta n.$$

De mesmo modo a variação de n é discreta de 1 em 1, resultando em:

$$\begin{aligned} s^2 &= \sum_{n=0}^N (n_i - \bar{x})^2 \cdot \frac{N!}{n!(N-n)!} \cdot p^n \cdot q^{N-n} \\ &= \sum_{n=0}^N (n^2 - 2n\bar{n} + \bar{n}^2) \cdot \frac{N!}{n!(N-n)!} \cdot p^n \cdot q^{N-n}. \end{aligned}$$

O que nos leva a analisar como a expressão de três somatórios: o primeiro é a média do quadrado de n

$$= \sum_{n=0}^N n^2 \cdot \frac{N!}{n!(N-n)!} \cdot p^n \cdot q^{N-n} = \overline{(n^2)};$$

O segundo é:

$$\begin{aligned} &= \sum_{n=0}^N -2n\bar{n} \cdot \frac{N!}{n!(N-n)!} \cdot p^n \cdot q^{N-n} \\ &= -2\bar{n} \cdot \overbrace{\sum_{n=0}^N n \cdot \frac{N!}{n!(N-n)!} \cdot p^n \cdot q^{N-n}}^{\bar{n}} \\ &= -2 \cdot \bar{n} \cdot \bar{n} \\ &= -2 \cdot (\bar{n})^2; \end{aligned}$$

o terceiro fica:

$$\begin{aligned} &= \sum_{n=0}^N (\bar{n})^2 \cdot \frac{N!}{n!(N-n)!} \cdot p^n \cdot q^{N-n} \\ &= (\bar{n})^2 \cdot \overbrace{\sum_{n=0}^N \frac{N!}{n!(N-n)!} \cdot p^n \cdot q^{N-n}}^1 \\ &= (\bar{n})^2. \end{aligned}$$

Portanto temos a variância da seguinte forma:

$$s^2 = \overline{(n^2)} - 2(\bar{n})^2 - (\bar{n})^2,$$

Resultando

$$s^2 = \overline{(n^2)} - (\bar{n})^2. \quad (3.3.12)$$

Agora utilizando novamente nossas variáveis a e b ,

$$a = N - 1 \text{ ou } N = a + 1$$

e

$$b = n - 1 \text{ ou } n = b + 1,$$

vamos calcular o valor de $\overline{n^2}$:

$$\begin{aligned} \overline{n^2} &= \sum_{n=0}^N n^2 \cdot \frac{N!}{n!(N-n)!} \cdot p^n \cdot q^{N-n} \\ \overline{n^2} &= \sum_{b=-1}^{a+1} (b+1)^2 \cdot \frac{(a+1)!}{(b+1)! \cdot (a-b)!} \cdot p^{b+1} \cdot q^{(a-b)}. \end{aligned}$$

Quando $b=-1$ o resultado é zero, portanto:

$$\begin{aligned} \overline{n^2} &= \sum_{b=0}^{a+1} (b+1)^2 \cdot \frac{(a+1)!}{(b+1)! \cdot (a-b)!} \cdot p^{b+1} \cdot q^{(a-b)} \\ \overline{n^2} &= \sum_{b=0}^{a+1} \frac{(b+1) \cdot (b+1) \cdot (a+1) \cdot a!}{(b+1) \cdot b! \cdot (a-b)!} \cdot p^{b+1} \cdot q^{(a-b)} \\ \overline{n^2} &= \sum_{b=0}^{a+1} \frac{(b+1) \cdot \overbrace{(a+1)}^N \cdot a!}{b! \cdot (a-b)!} \cdot p \cdot p^b \cdot q^{(a-b)} \\ \overline{n^2} &= N \cdot p \sum_{b=0}^{a+1} \frac{(b+1) \cdot a!}{b! \cdot (a-b)!} \cdot p^b \cdot q^{(a-b)}. \end{aligned}$$

Analisemos esta somatória da seguinte forma:

$$\overline{n^2} = N \cdot p \left[\sum_{b=0}^{a+1} \overbrace{b \cdot \frac{a!}{b! \cdot (a-b)!} \cdot p^b \cdot q^{(a-b)}}^{\overline{b=a \cdot p=(N-1) \cdot p}} + \underbrace{\sum_{b=0}^{a+1} \frac{a!}{b! \cdot (a-b)!} \cdot p^b \cdot q^{(a-b)}}_1 \right]$$

Assim temos:

$$\overline{n^2} = N \cdot p[(N-1) \cdot p + 1].$$

Voltando na equação na equação 3.3.12,

$$\begin{aligned}
s^2 &= \overline{(n^2)} - (\bar{n})^2 \\
s^2 &= Np[(N-1) \cdot p + 1] - (N \cdot P)^2 \\
s^2 &= Np[Np - p + 1] - N^2p^2 \\
s^2 &= N^2p^2 - Np^2 + Np - N^2p^2 \\
s^2 &= -Np^2 + Np = Np - Np^2 \\
s^2 &= Np[1 - p] \\
s^2 &= \bar{n} \overbrace{[1 - p]}^q,
\end{aligned}$$

assim obtemos o desvio padrão

$$s = \sqrt{N \cdot p \cdot q} = \sqrt{\bar{n} \cdot q}. \quad (3.3.13)$$

Assim, os dirigentes confirmaram que o jogador esteve apenas abaixo das expectativas.

3.3.5 Distribuição Normal

Pode se determinar uma densidade de probabilidade para um variável contínua à partir da distribuição binomial. Será demonstrado aqui um método retirado de [15], onde o erro será a diferença entre o valor observado e o valor verdadeiro

$$y = x - \bar{x}. \quad (3.3.14)$$

Vamos considerar as seguintes hipóteses:

- 1 Cada observação de uma grandeza que tem um valor verdadeiro está sujeita a erro.
- 2 Esse erro pode ser decomposto em um número muito grande de pequenos erros elementares. O número de fontes desses erros é sempre N , para todas as observações. O módulo do erro de qualquer das fontes é ε , igual para todas as fontes.
- 3 A probabilidade de um erro elementar ser positivo é igual à de ele ser negativo.

A partir disso, temos que responder qual é o desvio do valor real da grandeza quando das N fontes de erros, n derem erro para mais e as restantes $N-n$ derem erro para menos? O desvio será:

$$y = n\varepsilon + (N - n)(-\varepsilon) = (2n - N)\varepsilon. \quad (3.3.15)$$

Assim, qual é a probabilidade do erro ser y ? Essa probabilidade é dada pela função de probabilidade binomial com $p = q = \frac{1}{2}$

$$P_{N, \frac{1}{2}}(n) = \frac{N!}{(N - n)! n!} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^{N-n} = \frac{N!}{(N - n)! n!} \left(\frac{1}{2}\right)^N. \quad (3.3.16)$$

Como as probabilidades do erro ser para mais ou para menos foram supostas iguais, ambas tem de ser $\frac{1}{2}$. Assim a média do número de erros elementares para mais pode ser calculada pela fórmula 3.3.11.

$$\bar{n} = N \cdot p = \frac{1}{2}$$

Trazendo essa última equação na relação 3.3.15, verificamos que o erro y é nulo, como esperávamos pela terceira hipótese. O desvio padrão do número de erros pode ser mensurado pela fórmula 3.3.13:

$$s = \sqrt{N \cdot p \cdot q} = \frac{\sqrt{N}}{2}. \quad (3.3.17)$$

A tabela 3.3.6 relaciona as probabilidades do erro y com os números de erros elementares para mais ou para menos que causam o erro y . Para obter a função densidade de probabilidade para o erro, passamos a equação 3.3.16 ao limite que $N \rightarrow \infty$ e $\varepsilon \rightarrow 0$ para preservar a média e o desvio padrão do erro, grandezas consideradas nas hipóteses.

Observa-se que y e n são relacionados por um fator de escala, 2ε , com descolamento, $N\varepsilon$, constante. O descolamento $N\varepsilon$, por ser constante, não interfere na largura da função de probabilidade. O fator de escala deve ser considerado da seguinte forma, com a equação 3.3.17:

$$s_0 = \frac{\sqrt{N}}{2} \cdot 2\varepsilon = \sqrt{N}\varepsilon. \quad (3.3.18)$$

Onde s_0 é o desvio padrão de y . Como os erros são finitos, a função densidade de probabilidade de y tem um desvio padrão finito. Então, a restrição é que o limite da

Tabela 3.3.6: Relação entre as componentes elementares de erro e o erro resultante

Número de erros elementares		Erro Resultante	Probabilidade
Para mais	Para menos		
N	0	$N\varepsilon$	$\left(\frac{1}{2}\right)^N$
N-1	1	$(N-2)\varepsilon$	$N \cdot \left(\frac{1}{2}\right)^N$
N-2	2	$(N-4)\varepsilon$	$\frac{N(N-1)}{2} \left(\frac{1}{2}\right)^N$
...
n	N-n	$(2n-N)\varepsilon$	$\frac{N!}{(N-n)!n!} \left(\frac{1}{2}\right)^N$
...
0	N	$-N\varepsilon$	$\left(\frac{1}{2}\right)^N$

função 3.3.16 deve ser finito em $N\varepsilon^2$ apesar de $N \rightarrow \infty$ e $\varepsilon \rightarrow 0$. Antes de testar o limite, montemos a relação $\frac{\Delta P}{\Delta y}$ com ΔP como a variação da probabilidade entre obtermos $n+1$ erros *para mais* e n erros *para mais* e $\Delta y = 2\varepsilon$ pela equação 3.3.15, assim:

$$\frac{\Delta P}{\Delta y} = \frac{P(n+1) - P(n)}{2\varepsilon}$$

Pela aplicação de 3.3.16 retirada de [2], temos:

$$\begin{aligned} \frac{\Delta P}{\Delta y} &= \frac{1}{2\varepsilon} \cdot (P(n+1) - P(n)) \frac{1}{2\varepsilon} \\ &= \frac{1}{2\varepsilon} \left\{ \left[\frac{N!}{(n+1)!(N-n-1)!} \left(\frac{1}{2}\right)^N \right] - \left[\frac{N!}{n!(N-n)!} \left(\frac{1}{2}\right)^N \right] \right\} \\ &= \frac{1}{2\varepsilon} \left(\frac{1}{2}\right)^N \left\{ \left[\frac{N!}{(n+1)!(N-n-1)!} \right] - \left[\frac{N!}{n!(N-n)!} \right] \right\} \\ &= \frac{1}{2\varepsilon} \left(\frac{1}{2}\right)^N \left\{ \left[\frac{N!}{(n+1)n!(N-n-1)!} \right] - \left[\frac{N!}{n!(N-n)(N-n-1)!} \right] \right\} \\ &= \frac{1}{2\varepsilon} \cdot \frac{N!}{n!(N-n-1)!} \left(\frac{1}{2}\right)^N \cdot \left(\frac{1}{n+1} - \frac{1}{N-n} \right) \\ &= \frac{1}{2\varepsilon} \cdot \frac{N!}{n!(N-n-1)!} \left(\frac{1}{2}\right)^N \cdot \frac{N-n-(n+1)}{(n+1)(N-n)} \\ &= \frac{1}{2\varepsilon} \cdot \frac{N!}{n!(N-n)(N-n-1)!} \left(\frac{1}{2}\right)^N \cdot \frac{N-n-(n+1)}{(n+1)} \end{aligned}$$

$$= \frac{1}{2\varepsilon} \cdot \frac{N!}{n!(N-n)!} \left(\frac{1}{2}\right)^N \cdot \frac{N-n-(n+1)}{(n+1)}$$

obtendo-se

$$\frac{\Delta P}{\Delta y} = \frac{1}{2\varepsilon} P(n) \cdot \frac{N-2n-1}{n+1}$$

Temos N como um número muito grande, podemos desprezar o 1 em relação a n e a $N-2n$:

$$\frac{\Delta P}{\Delta y} = \frac{1}{2\varepsilon} P(n) \cdot \frac{N-2n}{n}$$

. Note que pela equação 3.3.15, temos que

$$N-2n = -\frac{y}{\varepsilon}$$

. Substituindo temos

$$\frac{\Delta P}{\Delta y} = \frac{1}{2\varepsilon} P(n) \cdot \left(-\frac{y}{n\varepsilon}\right)$$

$$\frac{\Delta P}{\Delta y} = -\frac{yP}{2n\varepsilon^2}$$

$$\frac{\Delta P}{P} = -\frac{\Delta y}{n} \cdot \frac{y}{2\varepsilon^2}$$

. Usando a equação 3.3.15 podemos substituir n

$$\frac{\Delta P}{P} = -\frac{\Delta y}{\frac{y}{2\varepsilon} + \frac{N}{2}} \cdot \frac{y}{2\varepsilon^2} = \frac{-y\Delta y}{y\varepsilon + N\varepsilon^2}$$

.

Agora, aplicando o limite de $\varepsilon \rightarrow 0$ e $N \rightarrow \infty$, com $N\varepsilon^2 = \sigma_0^2$, temos

$$\frac{dG}{G} = \lim_{\substack{N \rightarrow \infty \\ \varepsilon \rightarrow 0 \\ N\varepsilon^2 \rightarrow \sigma_0^2}} \left(\frac{\Delta P}{P}\right) = -\lim \left(\frac{y \cdot \Delta y}{\varepsilon y + N\varepsilon^2}\right) = -\frac{y dy}{\sigma_0^2}$$

.

P representa uma função discreta que, ao passarmos o limite, essa função torna-se contínua, $G(y)$. Assim podemos integrar os dois lados da equação:

$$\int \frac{1}{G} \cdot dG = \int \left[-\frac{y}{\sigma_0^2}\right] \cdot dy = -\frac{1}{\sigma_0^2} \cdot \int y dy$$

$$\ln G = -\frac{1}{\sigma_0^2} \cdot \frac{y^2}{2} = -\frac{y^2}{2\sigma_0^2}$$

$$G(y) = A \cdot e^{-\frac{y^2}{2\sigma_0^2}}$$

A constante A pode ser obtida pela condição de normalização das funções densidade de probabilidade

$$\int_{-\infty}^{\infty} G(y) \cdot dy = 1 \rightarrow \int_{-\infty}^{\infty} A \cdot e^{-\frac{y^2}{2\sigma_0^2}} dy = 1$$

Assim, chamando a integral de I

$$I = \int_{-\infty}^{\infty} A \cdot e^{-\frac{y^2}{2\sigma_0^2}} dy$$

como encontrado em [2], vamos elevar ao quadrado ambos os lados

$$I^2 = \left[\int_{-\infty}^{\infty} A \cdot e^{-\frac{y^2}{2\sigma_0^2}} dy \right]^2 = \left[\int_{-\infty}^{\infty} A \cdot e^{-\frac{y^2}{2\sigma_0^2}} dy \right] \cdot \left[\int_{-\infty}^{\infty} A \cdot e^{-\frac{y^2}{2\sigma_0^2}} dy \right]$$

$$I^2 = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} A \cdot e^{-\frac{y^2}{2\sigma_0^2}} A \cdot e^{-\frac{y^2}{2\sigma_0^2}} dy \right] dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A \cdot e^{-\frac{y^2}{2\sigma_0^2}} dx dy$$

Ainda em [2], encontramos a troca de variáveis: $\begin{cases} x^2 + y^2 = r^2 \\ dx dy = r dr d\theta \end{cases}$

$$I^2 = \int_0^{2\pi} \int_0^{\infty} A^2 \cdot e^{-\frac{r^2}{2\sigma_0^2}} r dr d\theta \begin{cases} u = \frac{-r^2}{2\sigma_0^2} \rightarrow \\ du = \frac{-2r dr}{2\sigma_0^2} \rightarrow r dr = -\frac{2\sigma_0^2 du}{2} \rightarrow r dr = -\sigma_0^2 du \\ r = 0 \rightarrow u = 0; r = \infty \rightarrow u = -\infty \end{cases}$$

$$I^2 = \int_0^{2\pi} \int_0^{\infty} A^2 \cdot e^u (-\sigma_0^2 du) d\theta = -A^2 \sigma_0^2 \int_0^{2\pi} \int_0^{\infty} e^u du d\theta$$

$$I^2 = -A^2 \sigma_0^2 \int_0^{2\pi} [e^u]_0^{-\infty} d\theta = -A^2 \sigma_0^2 \int_0^{2\pi} [e^{-\infty} - e^0] d\theta$$

$$I^2 = -A^2 \sigma_0^2 \int_0^{2\pi} [0 - 1] d\theta = A^2 \sigma_0^2 \int_0^{2\pi} d\theta = A^2 \sigma_0^2 [\theta]_0^{2\pi} = A^2 \sigma_0^2 [2\pi - 0]$$

$$I^2 = 2A^2\sigma_0^2\pi \rightarrow I = \sqrt{2A^2\sigma_0^2\pi} = 1$$

$$A\sigma_0\sqrt{2\pi} = 1 \rightarrow A = \frac{1}{\sigma_0\sqrt{2\pi}}$$

Recordando que $y = x - \bar{x}$, fica determinado a função densidade de probabilidade normal gaussiana:

$$G(x) = \frac{1}{\sigma_0\sqrt{2\pi}} \cdot e^{-\frac{(x-\bar{x})^2}{2\sigma_0^2}}$$

onde \bar{x} é o valor médio ou valor esperado e σ_0 é o desvio padrão.

ECONOMETRIA FINANCEIRA

4.1 SÉRIES TEMPORAIS

Em [14], temos que uma série temporal é qualquer conjunto de observações ordenadas no tempo. Por exemplo:

- Temperatura durante o dia/semana/mês.
- Quantidade de chuva em uma determinada região.
- Um registro das marés em um porto.
- As variações do Índice Bovespa.

De acordo com [14], quando estudamos uma série temporal podemos estar interessados em:

- Investigar o mecanismo gerador da série temporal. Analisando as alturas das ondas do porto de Santos, podemos querer saber como elas foram geradas.
- Fazer previsões de valores futuros. Em curto prazo, como séries de venda, produção, estoque, etc.; e em longo prazo, como séries populacionais, econômicas, etc.
- Descrever apenas o comportamento da série. Construção do gráfico, a verificação da existência de tendência, ciclos e variações sazonais, a construção de histogramas e diagramas de dispersão.
- Procurar periodicidades relevantes nos dados.

Neste trabalho vamos tentar encontrar periodicidades relevantes a partir dos retornos dos preços das ações da LINX. Vamos estudar o modelo ARIMA como forma de padronizar retornos que aparentemente não tem periodicidades relevantes.

4.2 DEFINIÇÕES INICIAIS

4.2.1 Retorno

Um dos objetivos em finanças é a avaliação de riscos de uma carteira de ativos (instrumentos) financeiros. Em [9] encontramos:

$$R_t = \frac{P_t - p_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1, \quad (4.2.1)$$

onde P_t é o preço do ativo no instante t (ano, mês, semana, dia, hora, minuto). Podemos escrever a equação 4.2.1 como

$$1 + R_t = \frac{P_t}{P_{t-1}}$$

R_t é usualmente chamado de taxa e retorno.

Denotando $p_t = \log P_t$ (sendo o logaritmo na base e), definimos o retorno composto continuamente ou simplesmente log-retorno como

$$r_t = \log \frac{P_t}{P_{t-1}} = \log(1 + R_t) = p_t - p_{t-1} \quad (4.2.2)$$

Podemos ainda obter, por regra de logaritmos:

$$R_t = e^{r_t} - 1 \quad (4.2.3)$$

Na prática é preferível trabalhar com retornos, que são livres de escala, do que com preços, pois retornos tem propriedades estatísticas mais interessantes como, estacionariedade (estudada em processos estocásticos) e ergodicidade. Nosso objetivo, com este trabalho, é modelar retornos. Diversas classes de modelos podem ser utilizadas para este fim, como: ARMA, ARCH, GARCH, também modelos de volatilidade estocástica. Essa parte mais densa, vamos estudar mais a frente.

Podemos definir também retorno multiperíodos. O retorno simples de período k , entre os instantes $t-k$ e t é:

$$R_t[k] = \frac{P_t - P_{t-k}}{P_{t-k}} \quad (4.2.4)$$

Em termos de retornos de um período podemos escrever 4.2.4 como:

$$\begin{aligned} 1 + R_t[k] &= (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}) \\ &= \frac{P_t}{P_{t-1}} \cdot \frac{P_{t-1}}{P_{t-2}} \cdots \frac{P_{t-k+1}}{P_{t-k}} = \frac{P_t}{P_{t-k}}, \end{aligned}$$

De modo que:

$$R_t[k] = \frac{P_t}{P_{t-k}} - 1 \quad (4.2.5)$$

Para facilitar comparações em horizontes diferentes é comum "anualizar" o retorno simples, considerando:

$$R_t[k]_{\text{anualizado}} = \left[\prod_{j=0}^{k-1} (1 + R_{t-j}) \right]^{\frac{1}{k}} - 1,$$

Que pode ser aproximado por:

$$\left(\frac{1}{k}\right) \sum_{j=0}^{k-1} R_{t-j}$$

Usando expansão de Taylor até primeira ordem.

Por sua vez, o log-retorno de período k fica:

$$r_t[k] = \log \frac{P_t}{P_{t-k}} = \log(1 + R_t[k]) = \sum_{j=0}^{k-1} \log(1 + R_{t-j}) = \sum_{j=0}^{k-1} r_{t-j}. \quad (4.2.6)$$

Por exemplo, uma semana corresponde a 5 dias de transações, de modo que o log-retorno continuamente composto em uma semana é:

$$r_t[5] = r_t + r_{t-1} + r_{t-2} + r_{t-3} + r_{t-4}$$

De qualquer modo, vale ressaltar que a expressão 4.2.6 é interessante do ponto de vista estatístico, pois para k relativamente grande a soma pode ser aproximada por um

valor amostral normal, usando o teorema do limite central que veremos mais adiante. Se houver pagamento de dividendos D_t no período, então os retornos ficam:

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1 \quad (4.2.7)$$

$$r_t = \log(1 + R_t) = \log(P_t + D_t) - \log P_{t-1} \quad (4.2.8)$$

Vemos que r_t é uma função não-linear de log-preços e log-dividendos.

Exemplo 4.1. Veja a seguir a diferença entre os gráficos das ações da Petrobrás PN (PETR4) quando visto em uma escala de preços normal (figura 4.2.1) e quando os preços são colocados em escala logarítmica (figura 4.2.2). Observe que o retorno

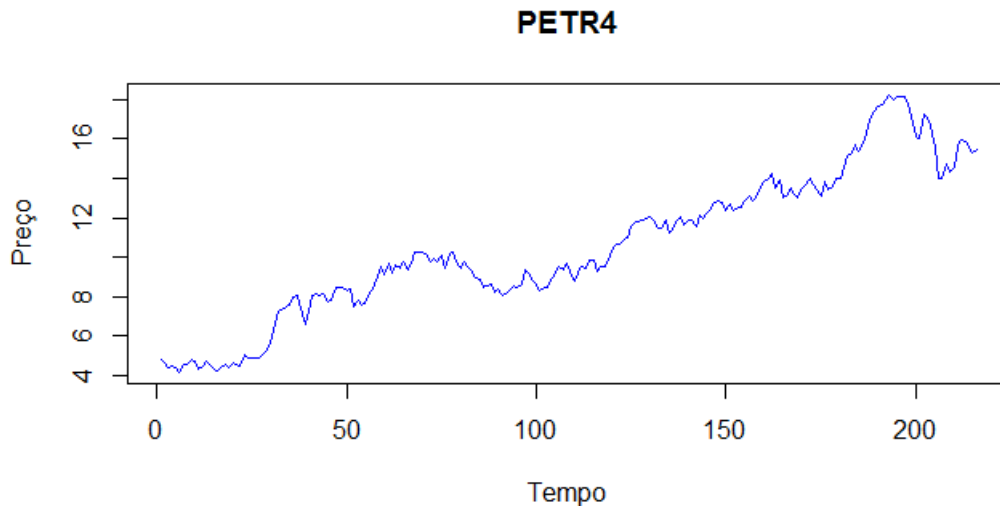


Figura 4.2.1: PETR4 em escala normal

entre os preços R\$ 4 e R\$ 8 é o mesmo retorno entre os preços R\$ 8 e R\$16. Porém essa distância no gráfico 4.2.1 não é analisada dessa forma, veja o gráfico na escala logarítmica 4.2.2. Agora o retorno entre os mesmos preços está proporcional. Veja que agora a distância entre R\$ 4 e R\$8 é a mesma entre R\$ 8 e R\$ 16.

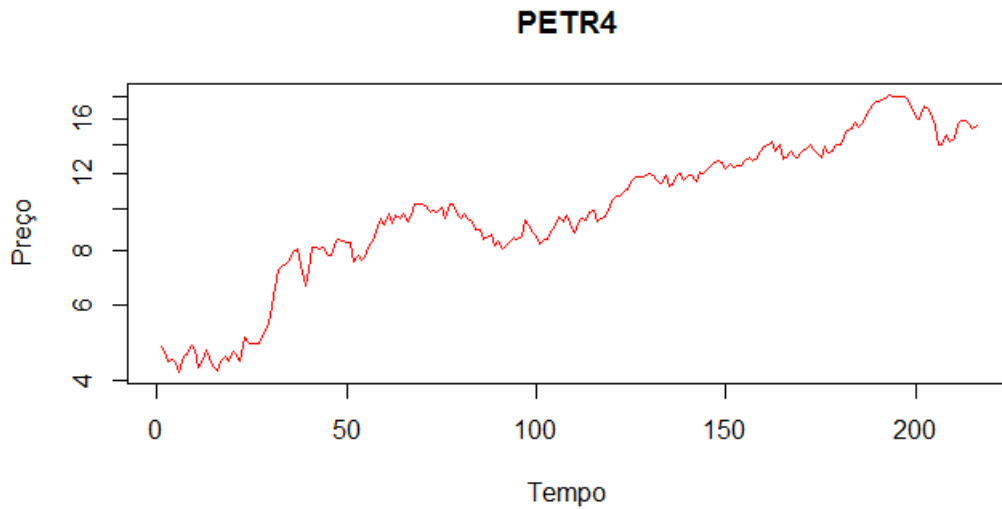


Figura 4.2.2: PETR4 em escala logarítmica

Agregação de Retornos

Na equação 4.2.6 temos o que se chama de agregação temporal dos retornos. Podemos ter também uma agregação “Cross-section”, para diversos ativos de uma carteira de investimentos c . Suponha que esta carteira contenha N instrumentos A_1, A_2, \dots, A_N , com pesos w_1, \dots, w_N , com $\sum_{i=1}^N w_i = 1$. Denotemos por R_i os retornos simples e por r_i os log-retornos desses ativos, $i = 1 \dots, N$. Se P_0 indicar o preço inicial da carteira, após um período teremos, para retornos continuamente compostos:

$$\frac{P_1}{P_0} = \sum_{i=1}^N w_i e^{r_i} \quad (4.2.9)$$

O log-retorno da carteira é $r_c = \log \frac{P_1}{P_0}$, logo obtemos:

$$r_c = \log \sum_{i=1}^N w_i e^{r_i} \quad (4.2.10)$$

Enquanto o retorno simples da carteira é:

$$\begin{aligned} R_c &= \frac{P_1}{P_0} - 1 \\ &= \sum_{i=1}^N w_i e^{r_i} - 1 \\ &= \sum_{i=1}^N w_i (1 + R_i) - 1. \end{aligned}$$

De outro modo:

$$R_c = \sum_{i=1}^N w_i R_i \quad (4.2.11)$$

No caso de composição discreta, temos:

$$\frac{P_1}{P_0} = \sum_{i=1}^N w_i (1 + r_i) \quad (4.2.12)$$

De modo que o retorno simples da carteira é $R_c = \frac{(P_1 - P_0)}{P_0}$, ou seja:

$$R_c = \sum_{i=1}^N w_i (1 + r_i) - 1 = \sum_{i=1}^N w_i R_i \quad (4.2.13)$$

Observe que nas equações 4.2.11 e 4.2.13, o retorno simples é uma soma ponderada de retornos simples, no caso de composição contínua, e uma soma ponderada de log-retornos, no caso de composição discreta. De modo geral:

- **Agregação temporal:** para $i = 1, \dots, N$

$$R_{i,t}[k] = \prod_{j=0}^{k-1} (1 + R_{i,t-j}) - 1 \quad (4.2.14)$$

$$r_{i,t} = \sum_{j=0}^{k-1} r_{i,j-t} \quad (4.2.15)$$

Para retornos simples e log-retornos, respectivamente.

- **Agregação "Cross-section":** para a carteira c e período t :

$$R_{ct} = \sum_{i=1}^N w_i R_{i,t} \quad (4.2.16)$$

$$r_{ct} = \log \left(\sum_{i=1}^N w_i e^{r_{it}} \right) \quad (4.2.17)$$

Para agregação temporal é mais conveniente trabalhar com log-retornos, enquanto que para agregação de ativos os retornos simples são mais convenientes. A expressão 4.2.17 é aproximada por:

$$r_{ct} \approx \sum_{i=1}^N w_i r_{i,t}$$

4.2.2 Especificações de um Processo Estocástico

Inicialmente, os processos estocásticos foram usados pelos físicos para descrever o movimento Browniano, que basicamente é o movimento de uma partícula de pólen imerso na em um copo com água. Estes processos podem ser divididos em: Processos Estacionários e Processos não-Estacionários. Para as variáveis financeiras, como preços, retornos e volatilidade o conceito de Processo Estocástico é usado para descrever seus comportamentos, tendência ou lateralização. A seguir algumas definições retiradas de [3].

Definição 7. (Espaço de índices ou Parâmetros temporais). O conjunto $T \neq \emptyset \subset \mathbb{R}$ é denominado espaço de parâmetros temporais ou índices do processo. O conjunto T possui uma ordem e vamos pensar que para cada $t \in T$ a variável X_t descreve o que acontece com o processo no tempo t em dois casos principais:

Parâmetro Temporal Discreto - $T = \mathbb{N}$, ou \mathbb{Z} , ou ainda $\{1, 2, \dots, n\}$

Parâmetro Temporal Contínuo - $T = [a, b]$, ou $T = \{t \in \mathbb{R} : t \leq 0\} = \mathbb{R}^+$ ou ainda $T = \mathbb{R}$

Definição 8. (Espaço de Estados). É o conjunto S dos estados possíveis de cada variável aleatória X_t .

Quando S é finito, será descrito na forma $\{1, 2, \dots, m\}$ onde m é um número natural. Caso S seja enumerável infinito, será escrito na forma $S = \mathbb{N}$, ou $S = \mathbb{Z}$.

Para todo $\omega \in \Omega$ fixado, e t também fixo, $X_t(\omega)$ determina o valor do processo no tempo t avaliado em ω e algumas vezes é escrito como w_t . Usaremos a letra grega ω para denotar $\omega = \{w_t\}_{t \in T} \in S^T$ associado a certo ω .

Em [9] entendemos que um processo estocástico é uma família de variáveis aleatórias que supomos definidas num mesmo espaço de probabilidades (Ω, \mathcal{F}, P) , onde Ω é um conjunto não-vazio (espaço amostral), \mathcal{F} é uma coleção de subconjuntos contidos em Ω e P é uma função que associa a cada elemento de \mathcal{F} um número real não-negativo (probabilidade associada ao evento). O conjunto T , quando tomado pelos conjuntos \mathbb{N} e \mathbb{Z} é denominado um processo de tempo discreto e quando tomado pelo conjunto dos reais \mathbb{R} ou $[0, \infty]$, é denominado um processo de tempo contínuo.

Definição 9. Seja T um conjunto arbitrário. Um processo estocástico é uma família $\{X_t, t \in T\}$, tal que, para cada $t \in T$, X_t é uma variável aleatória.

Sejam t_1, t_2, \dots, t_n elementos quaisquer de T e consideremos

$$F(x_1, \dots, x_n; t_1, \dots, t_n) = P\{X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n\}. \quad (4.2.18)$$

Então, o processo estocástico $\{X_t, t \in T\}$ estará especificado se conhecermos as distribuições *finito-dimensionais* - equação 4.2.18 - para todo $n \geq 1$.

As funções de distribuição da equação 4.2.18 devem satisfazer às duas condições seguintes:

- i- Condição de simetria: Para qualquer permutação j_1, j_2, \dots, j_n , dos índices $1, 2, \dots, n$ temos

$$F(x_{j_1}, \dots, x_{j_n}; t_{j_1}, \dots, t_{j_n}) = F(x_1, \dots, x_n; t_1, \dots, t_n) \quad (4.2.19)$$

- ii Condição de compatibilidade: para $m \leq n$,

$$F(x_1, \dots, x_m, x_{m+1}, \dots, x_n; t_1, \dots, t_m, \dots, t_n) = F(x_1, \dots, x_m; t_1, \dots, t_m) \quad (4.2.20)$$

Quando estudamos [9] percebemos que o conhecimento de todas as distribuições finito-dimensionais é muito difícil ocorrer na prática, senão impossível. O que se faz é estudar certas características associadas a 4.2.18 e que sejam simples de calcular e interpretar. Assim, os processos estocásticos que vão nos interessar são os processos estacionários de primeira e segunda ordem. Mas o que são processos estacionários?

4.3 PROCESSOS ESTOCÁSTICOS ESTACIONÁRIOS

Antes de continuar os estudos sobre processos estacionários vejamos duas relações. A *função média* X_t que será dada por:

$$\mu(1; t) = \mu_t = E[X_t] = \int_{-\infty}^{\infty} x dF(x; t) \quad (4.3.1)$$

e a função de autocovariância (f.a.c.v.) de X_t definida por:

$$\mu(1, 1; t_1, t_2) - \mu(1; t_1)\mu(1; t_2) = \gamma(t_1, t_2) = E[X_{t_1}X_{t_2}] - E[X_{t_1}]E[X_{t_2}] \quad (4.3.2)$$

Vemos que μ_t é uma função de $t \in T$ e que $\gamma(t_1, t_2)$ depende de dois argumentos. Em particular, se $t_1 = t_2 = t$, 4.3.2 fica:

$$\gamma(t, t) = \text{Var}[X_t] = E[X_t^2] - E^2[X_t] \quad (4.3.3)$$

onde 4.3.3 é a *função variância* do processo X_t e é indicada por σ_t^2

Definição 10. Um processo estocástico $X = \{X_t, t \in T\}$ é dito estritamente estacionário se todas as distribuições finito-dimensionais permanecem as mesmas sob translações no tempo:

$$F(x_1, \dots, x_n; t_1 + \tau, \dots, t_n + \tau) = F(x_1, \dots, x_n; t_1, \dots, t_n)$$

Definição 11. Um processo estocástico $\{X_t, t \in T\}$ chama-se estacionário de segunda ordem, se e somente se:

1. $E[X_t] = \mu$, constante $\forall t \in T$;
2. $E[X_t^2] < \infty$, $\forall t \in T$;
3. $\text{Cov}[X_t, X_{t-1}]$ não depende de t .

Um processo pode ser estritamente estacionário caso valha (1), mas só será estacionário de segunda ordem caso valha (2). Supõe-se μ (média) = 0, sem perda de generalidade. Caso contrário considere o processo $\{X_t - \mu\}$. De $\text{Cov}[X_t, X_{t-1}]$, podemos tirar que, para todo $t \geq 1$:

$$\text{Cov}[X_t, X_{t-\tau}] = \begin{cases} \sigma^2, & \text{se } \tau = 0 \\ 0, & \text{se } \tau \neq 0 \end{cases} \quad (4.3.4)$$

4.3.1 Propriedades da Função de Autocovariância

Dado um processo estacionário real discreto $\{X_t, t \in Z\}$, de média zero e função de autocovariância $\gamma_\tau = E[X_t X_{t+\tau}]$

Propriedade 6. A função de autocovariância γ_τ satisfaz as seguintes propriedades:

- i) $\gamma_\tau > 0$,
- ii) $\gamma_{-\tau} = \gamma_\tau$
- iii) $|\gamma_\tau| \leq \gamma_0$,

Em [9]

Demonstração. As propriedades (i) e (ii) decorrem imediatamente da definição de γ_τ . A propriedade (iii) segue do fato que:

$$E[X_{t+\tau} \pm X_t]^2 = E[X_{t+\tau}^2 \pm 2X_{t+\tau}X_t + X_t^2] \geq 0$$

Mas o segundo membro é igual a:

$$\sigma^2 \pm 2\gamma_\tau + \sigma^2 \geq 0$$

$$2\gamma_0 \pm 2\gamma_\tau \geq 0$$

E (iii) fica demonstrada. □

Com isso podemos definir, a partir de agora, a função de autocorrelação de um processo:

Definição 12 (Autocorrelação (f.a.c.)).

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0}, \tau \in Z$$

4.3.2 Funções de Autocorrelação e Autocorrelação Parcial:

Tanto uma quanto a outra são medidas que informam o quanto o valor de uma variável aleatória é capaz de influenciar seus vizinhos. Por exemplo, o quanto a existência de um valor mais alto condiciona valores também altos de seus vizinhos. Fato muito importante nos fechamentos dos preços.

Se o valor da autocorrelação está entre 1 (correlação perfeita) e -1, o que significa anti-correlação perfeita. O valor 0 significa total ausência de correlação. A autocorrelação de uma dada variável se define pela distância, ou atraso com que se deseja medi-la. Quando a FAC tende a zero, significa que não existe correlação entre os termos. A não existência de correlação implica em estacionariedade, caso contrário, não-estacionariedade. Para chegar à sua fórmula, usa-se a equação de autocovariância, denotada por γ_τ , dada abaixo:

$$\gamma_\tau = \text{cov}[X_t, X_{t-\tau}] = E[(X_t - \mu)(X_{t-\tau} - \mu)] = E[X_t X_{t-\tau}] - \mu^2. \quad (4.3.5)$$

No instante que $\tau=0$, γ_0 é igual à variância de X_t . Usando a equação 4.3.5, obtemos a função de autocorrelação, denotada por ρ_τ :

$$\text{FAC}_\tau = \rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \frac{\sum_{t=\tau+1}^T (X_t - \mu)(X_{t-\tau} - \mu)}{(T - \tau)\gamma_0} \quad (4.3.6)$$

Que tem as seguintes propriedades:

1. $\rho_0 = 1$
2. $\rho_\tau \in (-1, 1)$; $\tau \neq 0$

Função de Autocorrelação Parcial (FACP) é definida como o último coeficiente da autorregressão de ordem τ da série $Z_t = X_t - \mu$. Da mesma forma, quando a FACP tende a zero, significa que não existe correlação entre os termos. Devemos, de qualquer modo, ter em conta que as autocorrelações para intervalos sucessivos são formalmente dependentes. Pense-se no seguinte: se o primeiro elemento da série está fortemente relacionado com o segundo, e o segundo com o terceiro, então o primeiro está de alguma forma também relacionado com o terceiro. Daí que o coeficiente de autocorrelação para um atraso de ordem 2 não seja independente do coeficiente de autocorrelação para um atraso de ordem 1. Para obter uma informação sobre autocorrelações na série sem esta influência em cascata, definiu-se outro conceito, o de função de autocorrelação parcial. A autocorrelação parcial de atraso k corresponde à autocorrelação entre X_t e X_{t-k} que não é explicada pelos atrasos de 1 a τ . Esta medida corresponde à correlação de X_t e X_{t-k} removendo o efeito das observações $X_{t-1}, X_{t-2}, \dots, X_{t-k+1}$ e é denotada por ϕ_{kk} , ou seja

$$\phi_{kk} = \text{Corr}(X_t, X_{t-k} / X_{t-1}, X_{t-2}, \dots, X_{t-k+1})$$

Um método geral para encontrar a f.a.c.p. para um processo estacionário com f.a.c. ρ_k é utilizando as equações de relação matricial de Yule-Walker [9], isto é, para certo k temos:

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

Desenvolvendo a equação temos:

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

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

$$\vdots = \vdots$$

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

Resolvendo as equações acima sucessivamente para $k = 1, 2, \dots$, obtemos ϕ_{kk} . Em geral:

$$\phi_{kk} = \frac{|P_k^*|}{|P_k|}$$

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

Mais adiante vamos entender como usar a FAC e a FACP nos processos estacionários.

4.3.3 Processos Lineares Estacionários

Agora iniciaremos as definições dos processos auto-regressivos (AR), de médias móveis (MA) e a combinação destes, os processos chamados ARMA. Mais adiante, veremos como construir modelos ARMA, ou seja, como a partir dos dados podemos identificar um particular modelo que os represente. Em [9]:

Exemplo 4.2. Sequência Aleatória

Consideremos $\{X_n, n = 1, 2, \dots\}$ uma sequência de v.a. definidas no mesmo espaço amostral Ω . Aqui $T = \{1, 2, \dots\}$ e temos um processo com parâmetro discreto, ou uma sequência aleatória. Para todo $n \geq 1$, podemos escrever:

$$P[X_1 = a_1, \dots, X_n = a_n] =$$

$$= P[X_1 = a_1] \times P[X_2 = a_2 | X_1 = a_1] \times \cdots \times P[X_n = a_n | X_1 = a_1, \dots, X_{n-1} = a_{n-1}] \quad (4.3.7)$$

Em 4.3.7, os a_j representam estados do processo e o espaço dos estados pode ser tomado como conjunto dos reais. O caso mais simples é aquele em que temos uma sequência $\{X_n, n \geq 1\}$ de v.a. *mutuamente independentes* e neste caso 4.3.7 fica:

$$P[X_1 = a_1, \dots, X_n = a_n] = P[X_1 = a_1] \cdots P[X_n = a_n] \quad (4.3.8)$$

Se as v.a. X_1, X_2, \dots tiverem todas a mesma distribuição teremos uma sequência de v.a. independentes e identicamente distribuídas (i.i.d.). Neste caso, o processo X_n é estacionário. Se $E[X_n] = \mu$, $\text{Var}[X_n] = \sigma^2$, para todo $n \geq 1$, então:

$$\gamma_\tau = \text{Cov}[X_n, X_{n+\tau}] = \begin{cases} \sigma^2, & \text{se } \tau = 0 \\ 0, & \text{se } \tau \neq 0 \end{cases} \quad (4.3.9)$$

Segue que $\rho_\tau = 1$, para $\tau = 0$ e $\rho_\tau = 0$ caso contrário.

Definição 13. Dizemos que $\{\varepsilon_t, t \in Z\}$ é um ruído branco discreto se as v.a. ε_t são não correlacionadas, isto é, $\text{Cov}[\varepsilon_t, \varepsilon_s] = 0$, $t \neq s$

Se as variáveis aleatórias $\varepsilon_{t,s}$ são independentes, elas também serão não correlacionadas. Uma sequência de v.a. independentes e identicamente distribuídas, como definida acima, é chamada de *processo puramente aleatório*.

Ilustraremos a seguir na figura 4.3.1 a função de autocorrelação de um ruído branco denotado por $\varepsilon_t \sim \text{R.B.}(0, \sigma^2)$.

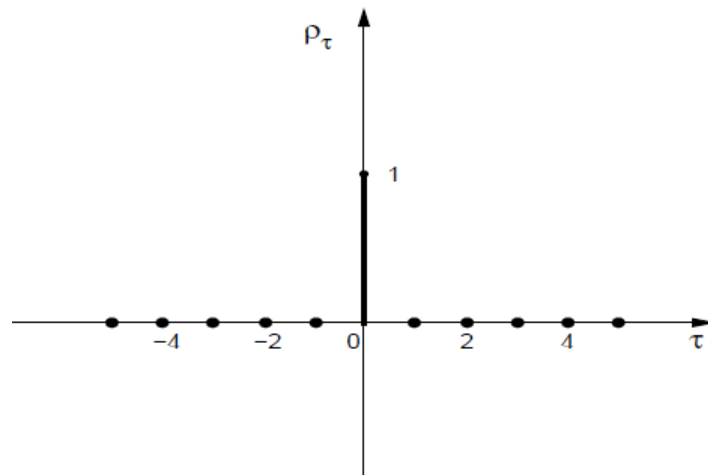


Figura 4.3.1: Função de autocorrelação de um ruído branco

4.3.4 Processos Auto-regressivos

Dizemos que $\{X_t, t \in Z\}$ é um processo auto-regressivo de ordem p , e escrevemos $X_t \sim \text{AR}(p)$, se satisfizer a equação de diferenças:

$$X_t - \mu = \phi(X_{t-1} - \mu) + \phi(X_{t-2} - \mu) + \dots + \phi(X_{t-p} - \mu) + \varepsilon_t$$

$$\mu, \phi_1, \dots, \phi_p \in \mathbb{R} \quad (4.3.10)$$

Mas $E[X_t] = \mu$, e se escrevermos o processo da seguinte forma:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

Então:

$$\mu = E[X_t] = \frac{\phi_0}{1 - \phi_1 - \dots - \phi_p}$$

Considerando agora o operador retroativo B como:

$$B^s X_t = X_{t-s}, s \geq 1$$

Então:

$$\phi(B)\tilde{X}_t = \varepsilon_t$$

Onde:

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

É o operador auto-regressivo de ordem p e \tilde{X} corresponde a valores subtraídos da média do processo:

$$\tilde{X}_t = X_t - \mu$$

Supondo $\mu = 0$, vejamos o processo AR(p) com $p = 1$:

$$X_t = \phi X_{t-1} + \varepsilon_t \quad (4.3.11)$$

Aqui temos que:

$$\phi(B) = 1 - \phi B$$

Com substituições sucessivas obtemos:

$$X_t = \sum_{j=0}^r \phi^j \varepsilon_{t-j} + \phi^{r+1} X_{t-r-1}$$

Se X_t for estacionário, com variância finita σ_X^2 , temos que:

$$E \left[X_t - \sum_{j=0}^r \phi^j \varepsilon_{t-j} \right]^2 = \phi^{2r+2} E[X_{t-r-1}^2] = \phi^{2r+2} \sigma_X^2$$

Agora, se $|\phi| < 1$, $\phi^{2(r+1)} \rightarrow 0$, quando $r \rightarrow \infty$, assim podemos escrever:

$$X_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j} \quad (4.3.12)$$

Onde a convergência é em média quadrática. Percebemos então que a condição $|\phi| < 1$ é suficiente para X_t ser estacionário. Multiplicando a equação 4.3.11 por $X_{t-\tau}$ e tomando a esperança, obtém-se:

$$\gamma_\tau = \phi\gamma_{\tau-1} = \dots = \phi^\tau\gamma_0$$

Mas de 4.3.12, obtemos:

$$\gamma_0 = \sigma_X^2 = \sigma^2 \sum_{j=0}^{\infty} \phi^{2j} = \frac{\sigma^2}{1 - \phi^2} \quad (4.3.13)$$

Assim se $\tau \geq 0$:

$$\gamma_\tau = \frac{\sigma^2}{1 - \phi^2} \phi^\tau$$

E como γ_τ é simétrica, podemos escrever a função de autocovariância de um processo AR(1) como:

$$\gamma_\tau = \frac{\sigma^2}{1 - \phi^2} \phi^{|\tau|}, \tau \in \mathbb{Z} \quad (4.3.14)$$

A função de autocorrelação de X_t é obtida da equação anterior e fica:

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \phi^{|\tau|}, \tau \in \mathbb{Z}. \quad (4.3.15)$$

Na próxima figura 4.3.3 temos as formas típicas de ρ_τ . Em ambos os casos, esta decai exponencialmente para zero, e este decai independente do sinal de ϕ .

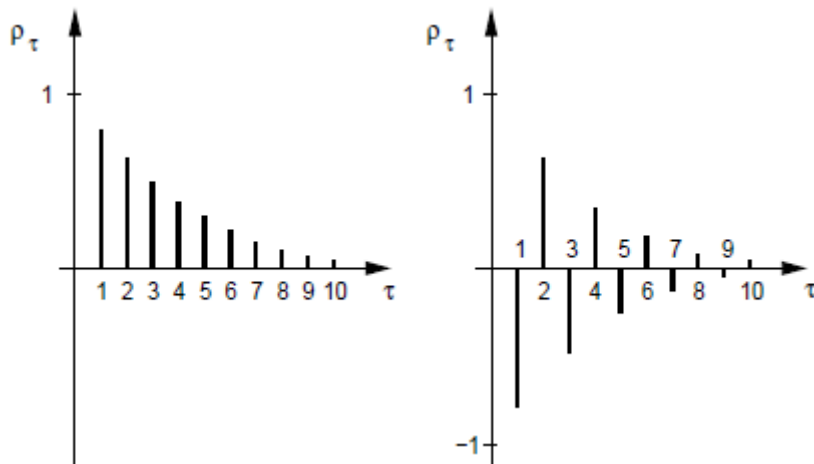


Figura 4.3.2: Função de autocorrelação de um processo AR(1) $\phi = 0,8$ e $\phi = -0,8$

Agora na figura 4.3.3 temos as funções de autocorrelação de um AR(2). Na primeira auto-regressão temos $\phi_1 = 0,5$ e $\phi_2 = 0,3$ e na segunda auto-regressão temos $\phi_1 = 1$ e $\phi_2 = -0,89$.

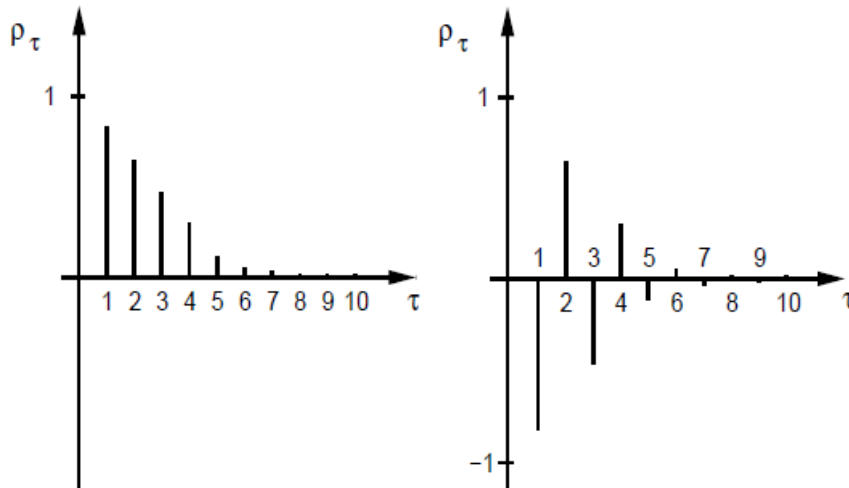


Figura 4.3.3: Função de autocorrelação de um processo AR(2) $\phi_1 = 0,5$ $\phi_2 = 0,3$ e $\phi_1 = 1$ $\phi_2 = -0,89$

4.3.5 Processos de Médias Móveis

Dizemos que $\{X_t, t \in Z\}$ é um processo de médias móveis de ordem q , denotado por $MA(q)$, se satisfizer à equação de diferenças:

$$X_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \tag{4.3.16}$$

Onde $\mu, \theta_1, \dots, \theta_q$ são constantes reais e $\varepsilon_t \sim R.B.(0, \sigma^2)$ Segue que X_t é estacionário, de média μ e como $\varepsilon_t \sim RB(0, \sigma^2)$

Se X_t for estacionário, de média μ e como o ε_t são não correlacionados, podemos obter a variância do processo como:

$$\sigma_X^2 = \sigma^2(1 + \theta_1^2 + \dots + \theta_q^2) \tag{4.3.17}$$

Suponha $\mu = 0$. e temos a função de autocovariância:

$$\begin{aligned} \gamma_\tau &= E\{X_T X_{t-\tau}\} = \\ \gamma(\tau) &= \sum_{k=1}^q \theta_k \gamma_\varepsilon(k - \tau) - \sum_{l=1}^q \theta_l \gamma_\varepsilon(\tau + l) + \sum_{k=1}^q \sum_{l=1}^q \theta_k \theta_l \gamma_\varepsilon(\tau + l - k) \end{aligned}$$

onde $\gamma_\varepsilon(\tau)$ é a função de autocovariância de ε_t . Logo:

$$\gamma_\tau = \begin{cases} \sigma^2(-\theta_\tau + \theta_1\theta_{\tau+1} + \dots + \theta_q\theta_{q-\tau}), & \text{se } \tau = 1, \dots, q \\ 0 & \text{se } \tau > q \\ \gamma - \tau & \text{se } \tau < q \end{cases} \quad (4.3.18)$$

Das equações 4.3.17 e 4.3.18 podemos obter a função de autocorrelação do processo MA(Q):

$$\rho_\tau = \begin{cases} \frac{-\theta_\tau + \theta_1\theta_{\tau+1} + \dots + \theta_q\theta_{q-\tau}}{1 + \theta_1^2 + \dots + \theta_q^2}, & \text{se } \tau = 1, \dots, q \\ 0 & \text{se } \tau > q \\ \rho - \tau & \text{se } \tau < q \end{cases} \quad (4.3.19)$$

Observa-se que ambos os processos se anulam para $|\tau| > q$. Em particular, para um processo MA(1):

$$X_t = \varepsilon_t - \theta\varepsilon_{t-1}$$

Obtém-se:

$$\text{Var}[X_t] = \sigma_X^2 = \sigma^2(1 + \theta^2),$$

$$\rho_\tau = \begin{cases} \frac{-\theta}{1 + \theta^2}, & \text{se } \tau = \pm 1 \\ 0, & \text{se } |\tau| > 1 \end{cases}$$

Definindo o operador de médias móveis de ordem q por:

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

A equação do processo de médias móveis 4.3.16, pode ser escrito:

$$X_t = \theta(B)\varepsilon_t$$

Em particular, para o processo MA(1) temos $\theta(B) = 1 - \theta B$, de modo que se pode escrever:

$$X_t = (1 - \theta B)\varepsilon_t$$

De onde segue que:

$$\varepsilon_t = (1 - \theta B)^{-1} X_t = (1 + \theta B + \theta^2 B^2 + \dots) X_t$$

Assim:

$$X_t = -\theta X_{t-1} - \theta^2 X_{t-2} - \dots + \varepsilon_t$$

Nesta última equação, temos que X_t é escrito como um processo auto-regressivo de ordem infinita. Dizemos que $|\theta| < 1$ é uma condição de invertibilidade para o processo MA(1). De modo geral, o processo de médias móveis 4.3.16 poderá ser escrito na forma:

$$X_t = \sum_{j=1}^{\infty} \pi_j X_{t-j} + \varepsilon_t$$

Se a condição de invertibilidade estiver satisfeita: todas as raízes de $\theta B = 0$ devem estar fora do círculo unitário.

4.3.6 O Processo ARMA(p,q)

Um processo ARMA é um processo composto pelos sistemas que acabamos de verificar AR(auto-regressivo) e MA(médias-móveis). Um ponto importante é como saber, olhando para uma série temporal, se ela segue um processo puramente AR e se sim, qual o valor de p? Ou um processo puramente MA e qual o valor de q? Ou diferente dos anteriores um processo misto, ARMA, e quais seriam o valor de p e q? Para isso vamos utilizar a Metodologia de Box-Jenkins (BJ) que de acordo com [14], consiste em identificar, ou seja, descobrir os valores apropriados para p e q, a partir da análise das funções de autocorrelação parcial e autocorrelação (FACP e FAC); estimar, ou seja, onde os parâmetros do modelo identificado são estimados e verificados. Também vamos diagnosticar através de uma análise de resíduos pode-se descobrir se o modelo é adequado. Caso o modelo não seja adequado, este ciclo de análise, estimação, verifi-

cação e diagnostico, é repetido. É importante destacar que para usar a metodologia de Box-Jenkins, deve-se ter, ou uma série temporal estacionária:

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \dots + \phi_p(X_{t-p} - \mu) + \theta_1\varepsilon_{t-1} - \dots - \theta_q\varepsilon_{t-q} \quad (4.3.20)$$

Como já vimos $\varepsilon_t \sim RB(0, \gamma^2)$. Segue que a média do processo é μ . Usando $\tilde{X}_t = X_t - \mu$ podemos escrever 4.3.20 na forma:

$$\phi(B)\tilde{X}_t = \theta(B)\varepsilon_t \quad (4.3.21)$$

Um modelo bastante usado é o ARMA(1,1). É escrito como:

$$X_t = \phi X_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1} \quad (4.3.22)$$

Para um processo ARMA(p,q) genérico a condição de estacionariedade é a mesma que para um processo AR(p), ou seja, as raízes de $\phi(B) = 0$ devem estar dentro do círculo unitário, e a condição de invertibilidade é a mesma para processos MA(q), ou seja, as raízes de $\theta(B) = 0$ devem estar dentro do círculo unitário.

Um fato interessante é que se tomarmos $\mu = 0$ e multiplicarmos a equação 4.3.20 por $X_{t-\tau}$ podemos obter a esperança da seguinte forma:

$$\gamma_\tau = \phi_1\gamma_{\tau-1} + \phi_2\gamma_{\tau-2} + \dots + \phi_p\gamma_{\tau-p} + \gamma X_{\varepsilon_\tau} - \theta_1\gamma X_{\varepsilon_{\tau-1}} - \dots - \theta_q\gamma X_{\varepsilon_{\tau-q}} \quad (4.3.23)$$

$\gamma X_{\varepsilon}(\tau)$ é a covariância entre X_t e ε_t escrita como:

$$\gamma X_{\varepsilon_\tau} = E[\varepsilon_t X_{t-\tau}]$$

Como $X_{t-\tau}$ só depende de choques ε_t ocorridos até o instante $t - \tau$ que esta covariância cruzada só é diferente de zero para $\tau \leq 0$, então:

$$\gamma_\tau = \phi_1\gamma_{\tau-1} + \phi_2\gamma_{\tau-2} + \dots + \phi_p\gamma_{\tau-p}, \quad \tau > q$$

Em [9] observamos que a conclusão é que as autocovariâncias de lags 1,2,...,q serão afetadas pelos parâmetros de médias móveis, mas para $\tau > q$, as mesmas se compor-

tam como nos modelos AR(p).

Para o caso de um modelo ARMA(1,1), obtemos:

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(1 - \phi\theta)(\phi - \theta)}{1 + \theta^2 - 2\phi\theta}$$

E, para $\tau > 1$

$$\rho_\tau = \phi\rho_{\tau-1}$$

Exemplo 4.3. Calcular a FAC e a FACP do modelo AR(1) e MA(1)

AR(1) Para AR(1) temos que $|\phi_1| < 1$

$$AR(1) = X_t = \phi_1 X_{t-1} + \varepsilon_t$$

$$E[X_t] = \mu = 0 \quad \text{Var}[X_t] = \gamma_0 \approx \sigma^2$$

Para uma melhor visualização, foi simulado um modelo AR(1) com $\phi_1 = 0,5$, com $\mu = 0$ e $\sigma^2 = 1$ como mostrado na figura 4.3.4.

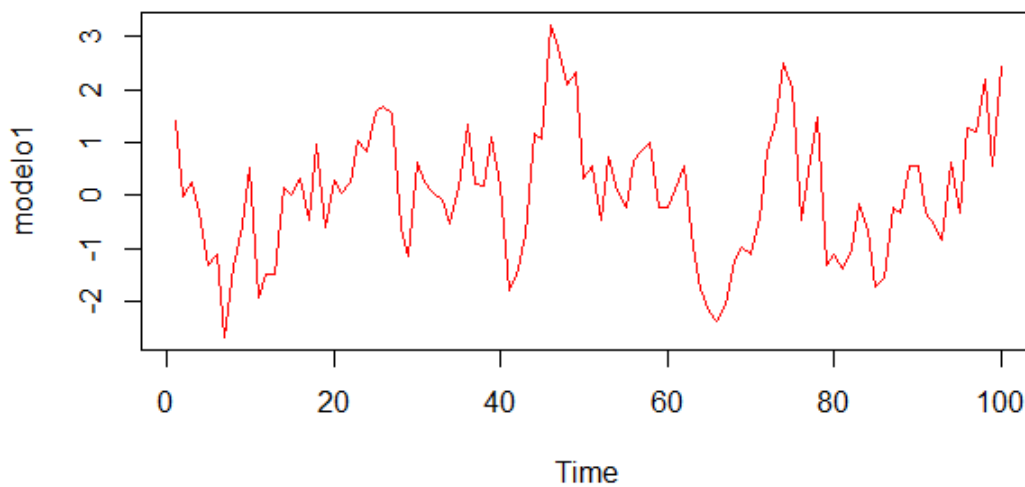


Figura 4.3.4: Modelo1: AR(1) simulado

Para $k \neq 0$, considera-se $Z_t = X_t - \mu = X_t$, pois $\mu = 0$:

$$X_{t+k} = \phi_1 X_{t+k-1} + \varepsilon_{t+k}$$

Multiplicando tudo por X_t :

$$X_t X_{t+k} = \phi_1 X_t X_{t+k-1} + X_t \varepsilon_{t+k}$$

Tomando a média e dividindo por γ_0 :

$$\rho_k = \phi_1 \rho_{k-1}$$

Por indução:

$$\rho_k = \phi_1 \rho_{k-1} = \phi_1^2 \rho_{k-2} = \dots = \phi_1^k \rho_0$$

portanto, $\text{fac}_k = \phi_1^k$. A partir da série simulada, percebe-se graficamente que o comportamento exponencial como mostra a figura 4.3.5.

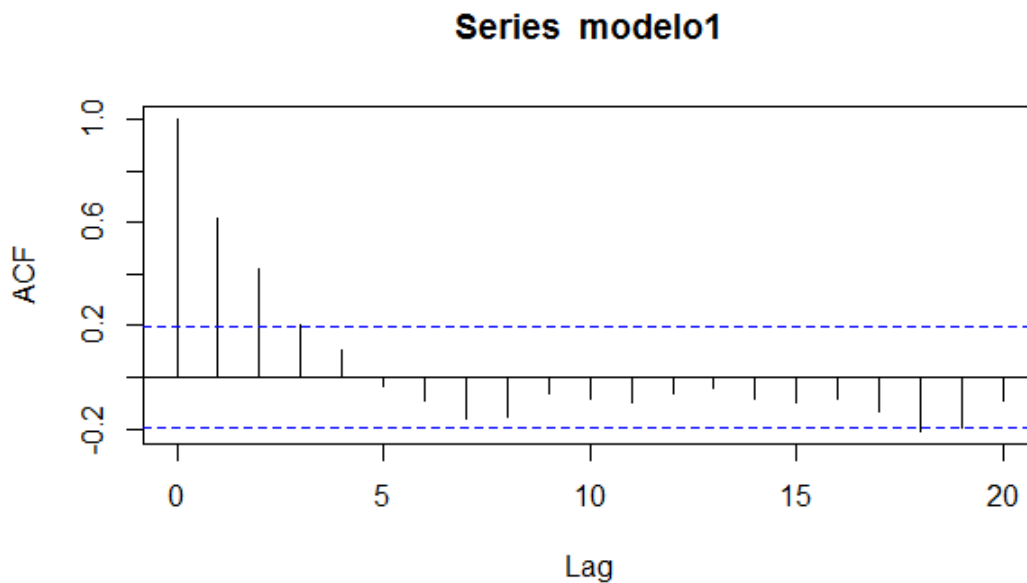


Figura 4.3.5: FAC do modelo 1

Como a FACP é o último coeficiente da autorregressão da série $Z_t = X_t - \mu$, mas $\mu = 0$, então $Z_t = X_t$, ou seja $\text{FACP} = \phi_1$. Como a ordem da série é $k=1$, só se considera a FACP nesse valor, isto é, caso $k > 1$ a $\text{FACP}_k = 0$. A figura 4.3.6 mostra a FACP da série simulada e a confirmação do comportamento esperado.

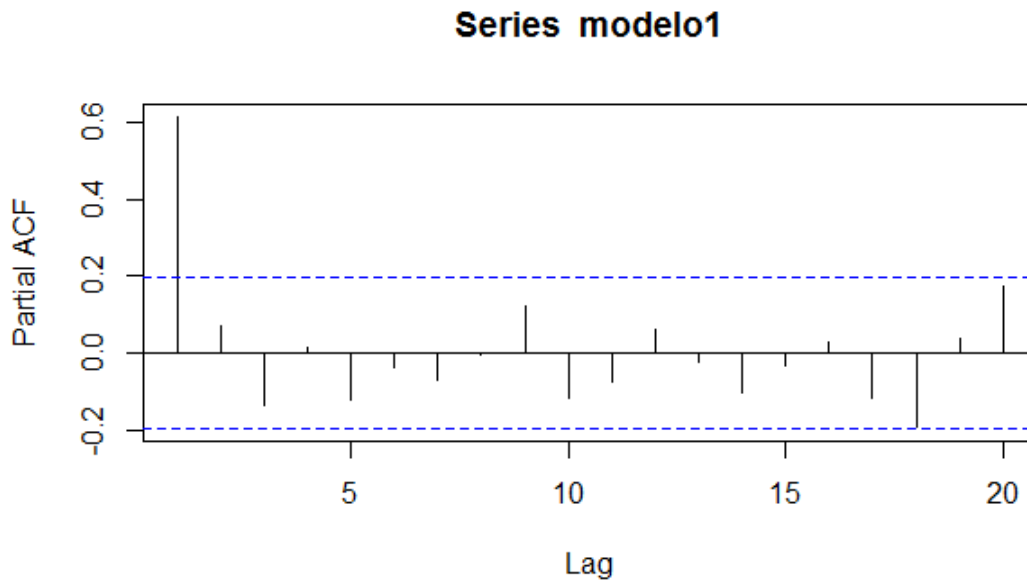


Figura 4.3.6: FACP do modelo 1

O mesmo ocorre quando $p > 1$, ou seja, na modelagem $AR(p)$ a $FACP_k = 0$, quando $k > p$ e $FAC_k \rightarrow 0$ exponencialmente, quando $k \rightarrow \infty$. Assim podemos perceber que o argumento p de $AR(p)$ é definido pela FACP.

MA(1) Para $MA(1)$, sabemos que $|\theta_1| < 1$:

$$MA(1) = X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

$$E[X_t] = \mu = 0$$

Assim como no exemplo anterior, vamos simular um modelo $MA(1)$ com $\theta_1 = 0.5$, $\mu = 0$ e $\sigma^2 = 1$, como mostrado na figura 4.3.7.

Primeiro multiplicamos a série por ela mesma e em seguida tiramos a média, de forma que se obtenha o valor de γ_0 :

$$X_t X_t = \theta_1 X_t \varepsilon_{t-1} + X_t \varepsilon_t$$

$$X_t X_t = \theta_1 [\theta_1 \varepsilon_{t-1} + \varepsilon_t] \varepsilon_{t-1} + [\theta_1 \varepsilon_{t-1} + \varepsilon_t] \varepsilon_t$$

Tirando a média:

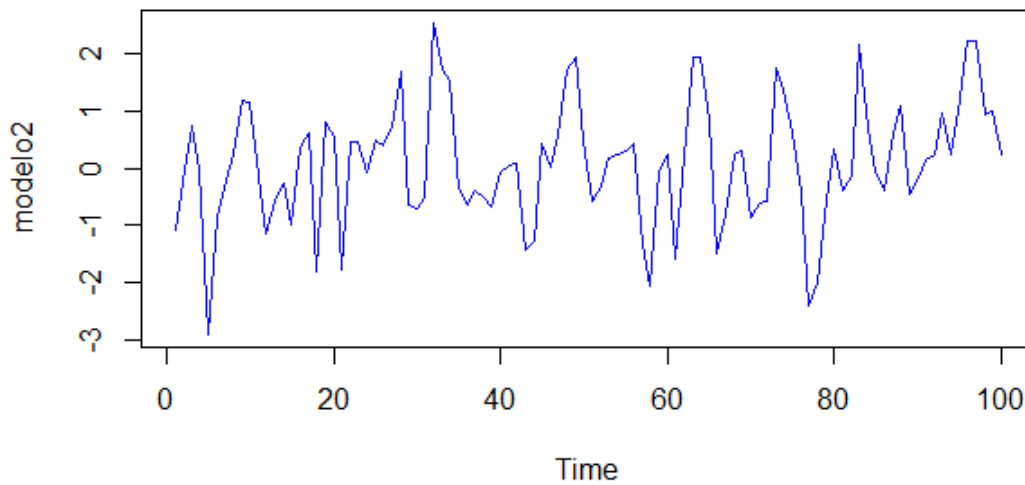


Figura 4.3.7: Modelo2: MA(1) simulado

$$\gamma_0 = \sigma^2(\theta_1^2 + 1)$$

Agora vamos multiplicar X_t por X_{t-1} , em seguida calculamos a média de forma que obtemos γ_1 :

$$X_{t-1}X_t = \theta_1 X_{t-1}\varepsilon_{t-1} + X_{t-1}\varepsilon_t$$

$$X_t X_t = \theta_1 [\theta_1 \varepsilon_{t-2} + \varepsilon_{t-1}] \varepsilon_{t-1} + [\theta_1 \varepsilon_{t-2} + \varepsilon_{t-1}] \varepsilon_t$$

Tirando a média:

$$\gamma_1 = \theta_1 \sigma^2$$

A FAC_1 será:

$$p_1 = \frac{\gamma_1}{\gamma_0} = \frac{\theta_1}{\theta_1^2 + 1}$$

Refazendo o mesmo procedimento, mas multiplicando X_t por X_{t-k} temos:

$$\gamma_k = 0$$

Portanto, $FAC_k = p_k = 0$, quando $k > 1$ como podemos ver na figura 4.3.8.

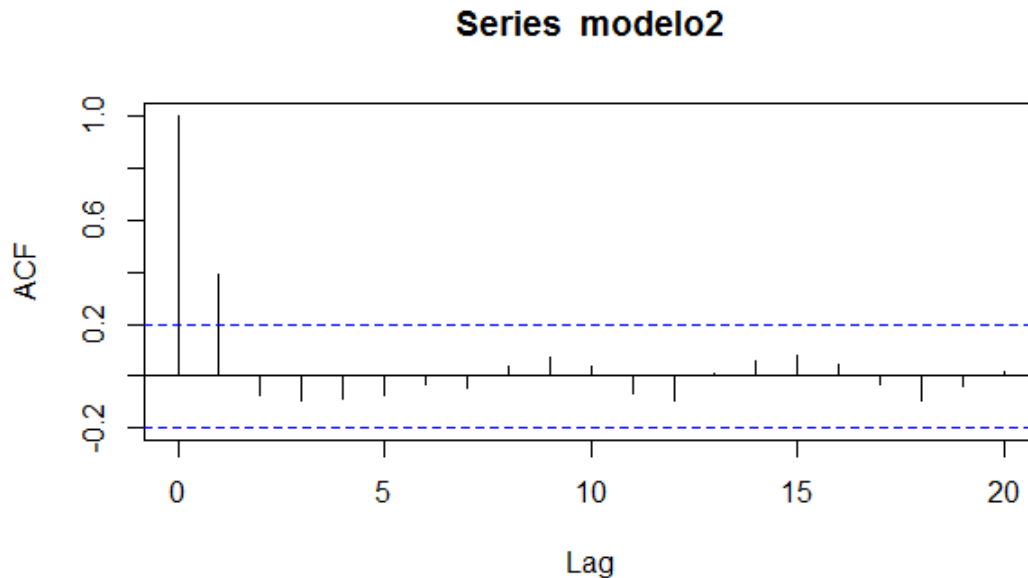


Figura 4.3.8: ACF do modelo 2

Para calcular a FACP usa-se as equações de Yule-Waller, presentes em [14], e é possível mostrar que a $FACP_k$ é uma exponencial. Para a visualização de tais acontecimentos foi gerado o gráfico da figura 4.3.9. É observado o mesmo comportamento para $q > 1$, ou seja, na modelagem $MA(q)$ a $FAC_k = 0$ quando $k > q$ e $FACP_k \rightarrow 0$ quando $k \rightarrow \infty$. Isso mostra que o argumento de $MA(q)$ é definido pela FAC.

4.4 PROCESSOS ESTOCÁSTICOS NÃO ESTACIONÁRIOS

Os processos que não apresentam médias ou variâncias constantes são chamados Não Estacionários. Será um Processo Não Estacionário Homogêneo se, ao tomarmos suas diferenças sucessivas, encontramos um processo estacionário.

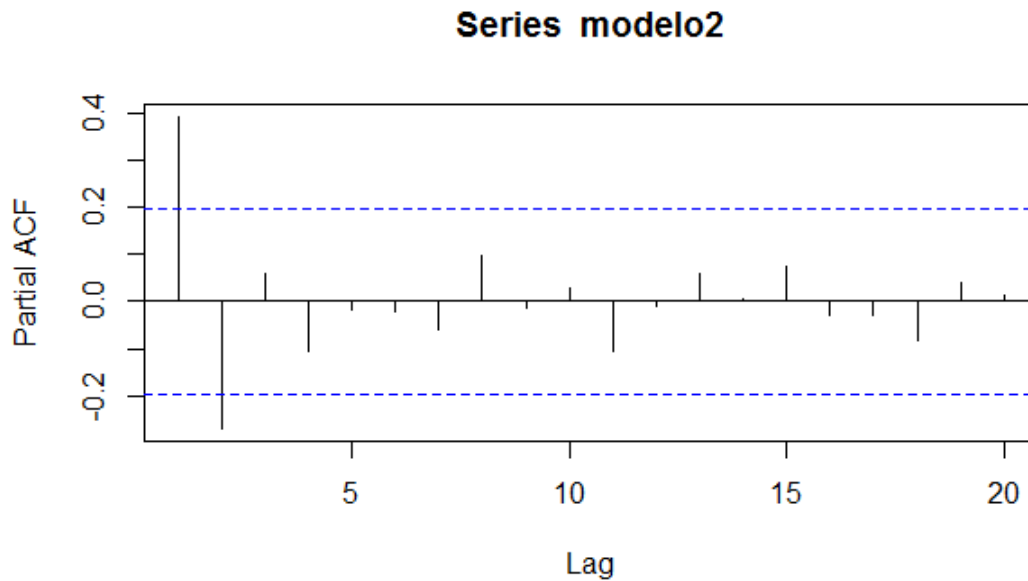


Figura 4.3.9: FACP do modelo 2

4.4.1 *Passeio Aleatório*

O passeio aleatório simples em \mathbb{Z} é um processo estocástico que descreve o movimento de uma partícula cuja posição no tempo t é dada pela variável aleatória X_t . A partícula se move nas vizinhanças de \mathbb{Z} e pode ir de um ponto x para um de seus vizinhos $x+1$ ou $x-1$, tendo probabilidade p ou $q = p - 1$ respectivamente, com $p \in [0,1]$. A sequência do p.a. é definida por $(X_t)_{t \geq 0}$. Essas probabilidades independem do ponto x . É chamado p.a. simétrico se $p, q = \frac{1}{2}$

Recorrência/Transiência

Supondo que uma partícula comece seu movimento na origem $X_0 = 0$, é interessante saber se, após iniciado o passeio, ela voltará à origem, após um intervalo de tempo tendendo ao infinito. Quando a partícula volte ao ponto inicial dizemos que o intervalo é finito e definiremos esse tempo como:

$$T_0 = \inf\{n > 0 : X_n = 0\}$$

Caso a trajetória não volte nunca à origem, o intervalo de tempo é infinito denotado por:

$$T_0 := \infty$$

E assim, se $P(T_0 < \infty) = 1$, o p.a. é dito recorrente. No contrário $P(T_0 < \infty) < 1$ o p.a. é dito transiente.

Em [12] encontramos que é importante saber que:

$$P(X_t = 0 | X_0 = 0) = \mu_t$$

temos $\mu_t = 0$, se t é ímpar, e que:

$$P(X_{2t} = 0 | X_0 = 0) = \mu_{2t} = \binom{2t}{n} p^t (1-p)^t$$

No decorrer do caminho percorrido, a partícula pode se mover para direita com probabilidade p , e para esquerda com probabilidade $q=1-p$. Assim, no espaço $2n$ a probabilidade é:

$$p^t (1-p)^t$$

Seja Y uma variável aleatória que denota o número de vezes, dentre os $2t$ primeiros passos, que a partícula se desloca para a direita. Como $Y \sim B(2t, p)$ e $P(X_{2t} = 0 | X_0 = 0) = P(Y = t)$, temos:

$$\mu_{2t} = \frac{(2t)!}{(t!)^2} p^t (1-p)^t$$

Usando a fórmula de Stirling (ver [12]) tem-se que para $p \neq \frac{1}{2}$ o p.a. é transiente e para $p = \frac{1}{2}$ o passeio é recorrente.

4.4.2 Movimento Browniano

A ideia do Movimento Browniano veio com as tentativas do biólogo Robert Brown de descrever o movimento de uma partícula de pólen suspensa na água.

Definição 14. Chamaremos de Movimento Browniano o processo contínuo $\{W_t, t > 0\}$ tal que:

- (a) $W(0)=0$;
- (b) Para quaisquer instantes $0 < t_1 < t_2 < \dots < t_k < 1$, as v.a.s $W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2}, \dots, W_{t_k} - W_{t_{k-1}}$ são independentes;
- (c) para quaisquer s, t e τ não negativos, as v.a. $W_t - W_s$ e $W_{t-\tau} - W_{s-\tau}$ tem a mesma distribuição;
- (d) para todo $t > 0$, W_t tem distribuição $N(0, t)$;
- (e) as trajetórias de W_t são contínuas com probabilidade um.

Observe que o movimento browniano tem incrementos independentes e estacionários. Como as distribuições finito-dimensionais de W são normais, então o processo é gaussiano. E finalmente, como:

$$W_t - W_s = W_{t-s} - W_0 = W_{t-s}$$

então, $W_t - W_s$ e W_{t-s} possuem distribuição $N(0, t - s)$

por definição do movimento browniano, $E[W_t] = 0$ e $\text{Var}[W_t = t]$, $t > 0$.

Em [9] encontramos que as trajetórias W são contínuas e irregulares devido aos seus incrementos serem independentes.

Na imagem 4.4.1, 3 movimentos iniciados em uma ação com um preço a 10R\$, com um $\sigma = 0.25$ e $\mu = 0.01$. Ver a programação em Python no Apêndice B.

4.4.3 Processo ARIMA(p, d, q)

Os modelos listados na seção 4.3 são apropriados para descrever séries estacionárias, isto é, séries que se desenvolvem no tempo ao redor de uma media constante. Porém as séries econômicas não são sempre estacionárias e para predizê-las necessitamos de um modelo que comporte a não estacionariedade. Um fato interessante é que mesmo não sendo estacionárias, algumas séries econômicas se tornam estacionárias quando diferenciadas [14]. Por exemplo Z_t é não estacionária, mas:

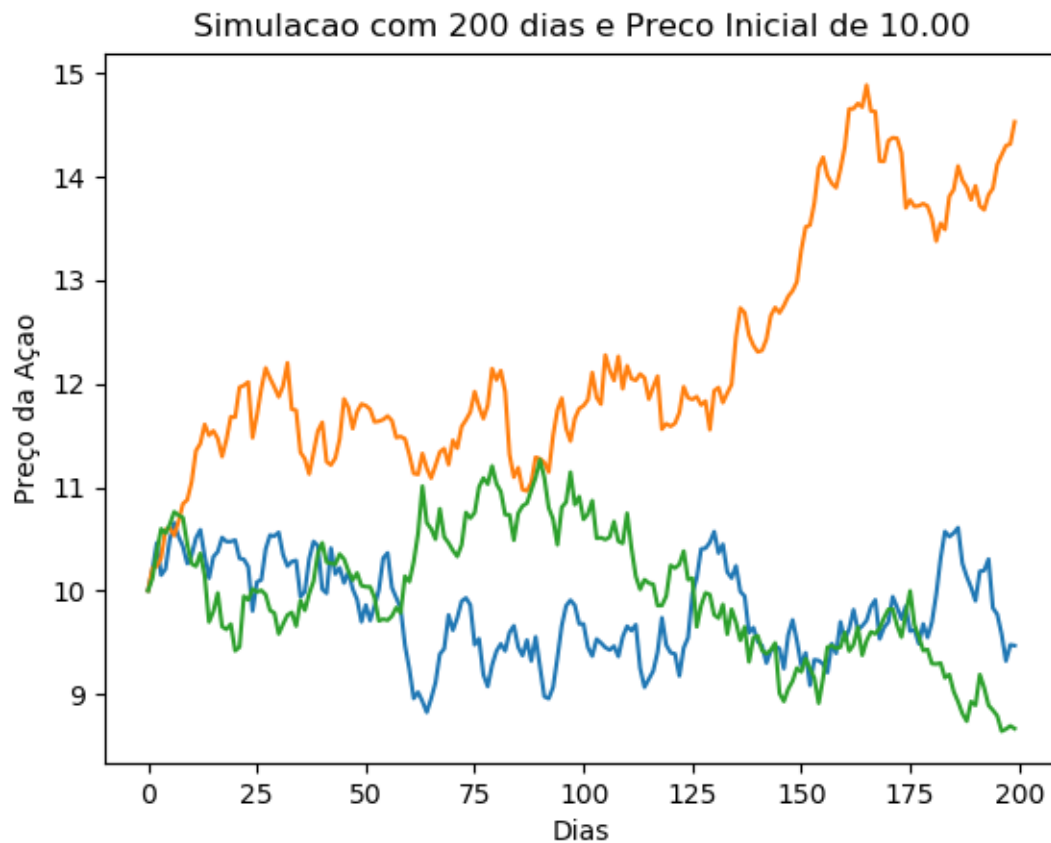


Figura 4.4.1: Movimento Browniano gerado em Python

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

é estacionária.

Portanto, se for preciso diferenciar uma série temporal d vezes para torná-la estacionária e então aplicar o modelo ARMA(p,q), diz-se que a série temporal original é ARIMA(p,d,q), ou seja, é uma série temporal autorregressiva integrada de média móvel, em que p indica o número de termos autorregressivos, d o número de vezes em que a série tem que ser integrada para se tornar estacionária e q o número de termos

de média móvel. Uma série pode apresentar várias formas de não estacionariedade. Considere um modelo AR(1):

$$(1 - \phi B)\bar{Z}_t = \alpha_t \quad (4.4.1)$$

Como já vimos, para que haja estacionariedade é necessário que $|\phi| < 1$. Quando $|\phi| = 1$ obtemos um processo não estacionário $\bar{Z}_t = \bar{Z}_{t-1} + \alpha_t$ e quando $|\phi| > 1$ o processo 4.4.1, assim como o processo 4.3.11 explode na medida em que t aumenta.

Vamos tratar aqui de séries cujo comportamento é não explosivo. Vamos estudar em particular as séries que apresentam alguma homogeneidade em seu comportamento. As séries Z_t tais que, tomando-se um número finito de diferenças, d , tornam-se estacionárias, são chamadas, como já vimos, não estacionárias homogêneas.

Se:

$$w_t = \Delta^d Z_t$$

É estacionária, podemos representar W_t por um modelo ARMA(p, q), ou seja:

$$\phi(B)W_t = \theta(B)\alpha_t \quad (4.4.2)$$

Se W_t é uma diferença de Z_t , então Z_t é integral de W_t , daí dizemos que Z_t segue modelo auto-regressivo-integrado-médias móveis, modelo ARIMA (p, d, q):

$$\phi(B)\Delta^d Z_t = \theta(B)\alpha_t \quad (4.4.3)$$

Os parâmetros p e q são as ordens de $\phi(B)$ e $\theta(B)$ respectivamente.

No modelo 4.4.2 todas as raízes de $\phi(B)$ estão fora do círculo unitário. Escrever 4.4.3 é equivalente a escrever:

$$\xi(B)Z_t = \theta(B)\alpha_t \quad (4.4.4)$$

Onde $\xi(B)$ é operador auto-regressivo não estacionário, de ordem $p+d$, com d raízes iguais a um e as restantes p fora do círculo unitário, ou seja:

$$\xi(B) = \phi(B)\Delta^d = \phi(B)(1 - B)^d$$

Assim, o modelo 4.4.3 supõe que a d -ésima diferença da série pode ser representada por um modelo ARIMA estacionário e invertível. Na maioria dos casos usuais uma ou duas diferenciações são suficientes. Veja dois casos interessantes e comuns de não estacionariedade homogênea [14]:

- i) séries não estacionárias quanto ao nível: oscilam ao redor de um nível médio durante algum tempo e depois saltam para outro nível temporário. Para torná-las estacionárias é suficiente tomar uma diferença; caso típico das séries econômicas;
- ii) séries não estacionárias quanto à inclinação: oscilam numa direção por algum tempo e depois mudam para outra direção temporária. Para torná-las estacionárias é necessária uma segunda diferença.

4.5 CONSTRUÇÃO DE MODELOS ARIMA

Em [14] encontramos os ciclos iterativos do método de Box & Jenkins, são eles:

- a) identificação
- b) estimação
- c) verificação

A fase mais crítica do método é a identificação dos índices p , d e q do modelo ARIMA. Está colheita é feita com base nas funções de autocovariância e autocorrelação parcial.

Lembramos que a equação função de autocorrelação é dada por:

$$r_j = \frac{c_j}{c_0} \quad (4.5.1)$$

Onde c_j é a estimativa da facv

$$c_j = \frac{1}{N} \sum_{t=1}^{N-j} [(Z_t - \bar{Z})(Z_{t+j} - \bar{Z})] \quad (4.5.2)$$

Sendo $\bar{Z} = \frac{1}{N} \sum_{t=1}^N Z_t$ a média amostral.

A variância de r_j , em um processo estacionário normal, pode ser aproximada por:

$$\text{Var}(r_j) = \frac{1}{N} \sum_{v=-\infty}^{\infty} [p_v^2 + p_{v+j}p_{v-j} - 4p_j p_v p_{v-j} + 2p_v^2 p_j^2] \quad (4.5.3)$$

Quando temos um processo em que as auto-correlações são nulas para $v > q$, todos os termos de 4.5.3 anulam-se para $j > q$, exceto o primeiro e obtemos:

$$\text{Var}(r_j) = \frac{1}{N} \left[1 + 2 \sum_{v=1}^q p_v^2 \right] \quad (4.5.4)$$

Como não conhecemos as auto-correlações p_v , substituímos por r_v , obtendo-se uma estimativa para 4.5.4:

$$\sigma^2(r_j) = \frac{1}{N} \left[1 + 2 \sum_{v=1}^q r_v^2 \right] \quad (4.5.5)$$

Para facp vimos que sob a hipótese que o processo é AR(p):

$$\text{Var}(\phi_{jj}) = \frac{1}{N} \quad (4.5.6)$$

De modo que:

$$\sigma(\phi_{jj}) = \frac{1}{\sqrt{N}} \quad (4.5.7)$$

O motivo de ser fazer a identificação é determinar os valores de p, d e q do modelo ARIMA(p, d, q). Para a identificação usaremos os passos de [14] que são:

- a) Faremos diferenciações da série Z_t , tantas vezes quantas necessárias a fim de se obter uma série estacionária. O processo $\Delta^d Z_t$ seja reduzido a um ARMA(p, q). O número d de diferenças necessárias para que o processo se torne estacionário é alcançado com a fac amostral de $W_t = \Delta^d Z_t$ decresce rapidamente pra zero;
- b) identificamos o processo ARMA(p, q) resultante através da análise das auto-correlações e auto-correlações parciais estimadas, cujos comportamentos devem imitar os comportamentos das respectivas quantidades teóricas.

A justificativa do item a) acima é que para um modelo ARMA estacionário, as fac são dadas por:

$$p_j = A_1 G_1^j + A_2 G_2^j + A_p G_p^j \quad (4.5.8)$$

Como $\phi(B) = \prod_{i=1}^p (1 - G_i B)$ onde as raízes de $\phi(B)$ devem estar fora do círculo unitário, devemos ter $|G_i| < 1$ Segue de 4.5.8 que se nenhuma raiz está muito próxima do círculo unitário, as auto-correlações p_j decairão para zero, para valores moderados de j .

Tabela 4.5.1: Tabela para ordens (1,d,0) e (0,d,1)

Ordem	(1,d,0)	(0,d,1)
Comportamento p_k	decai exponencialmente	somente $p_1 \neq 0$
Comportamento ϕ_{kk}	somente $\phi_{11} \neq 1$	decaimento exponencial dominante
Estimativas iniciais	$\phi = p_1$	$p_1 = \frac{-\theta_1}{1+\theta_1^2}$
Região de Admissão	$-1 < \phi_1 < 1$	$-1 < \theta_1 < 1$

Tabela 4.5.2: Tabela para ordens (2,d,0) e (0,d,2)

Ordem	(2,d,0)	(0,d,2)
Comportamento p_k	mistura de exponenciais ou ondas senoides amortecidas	somente $p_1 \neq 0$ e $p_2 \neq 0$
Comportamento ϕ_{kk}	somente $\phi_{11} \neq 1$ e $\phi_{22} \neq 0$	dominada por mistura de exponenciais ou senoides amortecidas
Estimativas iniciais	$\phi_1 = \frac{p_1(1-p_2)}{1-p_1^2}$ $\phi_2 = \frac{p_2-p_1^2}{1-p_1^2}$	$p_1 = \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2}$ $p_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}$
Região de Admissão	$-1 < \phi_2 < 1$ $\phi_2 - \phi_1 < 1$ $\phi_2 + \phi_1 < 1$	$-1 < \theta_2 < 1$ $\theta_2 - \theta_1 < 1$ $\theta_2 + \theta_1 < 1$

Porém se uma raiz real, G_1 , estiver próxima de um, $G_1 = 1 - \epsilon$, $\epsilon > 0$ pequeno. Como $G_1^j = (1 - \epsilon)^j \simeq 1 - j\epsilon$, vem que $p_j = A_1(1 - j\epsilon)$, o que mostra que a fac decairá lentamente para zero e de forma quase linear.

Na prática, $d=0,1$ ou 2 é o suficiente para inspecionar as primeiras 15 ou 20 autocorrelações da série e de suas diferenças.

Convém testar se $E(W_t) = \mu_w$ é zero, comparando \bar{W} com seu desvio padrão estimado. Retiramos de [14] uma tabela com o comportamento das fac e facp de um processo ARIMA(p,d,q).

O programa para identificação, estimação e verificação de modelos da família ARIMA foi o R STUDIO e a série utilizada foi o fechamento das ações da Linx (linx3) no período de 04-11-2015 até 15-12-2015, com 30 observações. As saídas serão:

Tabela 4.5.3: Tabela para ordens (1,d,1)

Ordem	(1,d,1)
Comportamento p_k	decai exponencialmente após lag 1
Comportamento ϕ_{kk}	dominada por decaimento exponencial após o lag 1
Estimativas iniciais	$p_1 = \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1-\theta_1^2-2\phi_1\theta_1}$ $p_2 = p_1\theta_1$
Região de Admissão	$-1 < \phi_1 < 1$ $-1 < \theta_1 < 1$

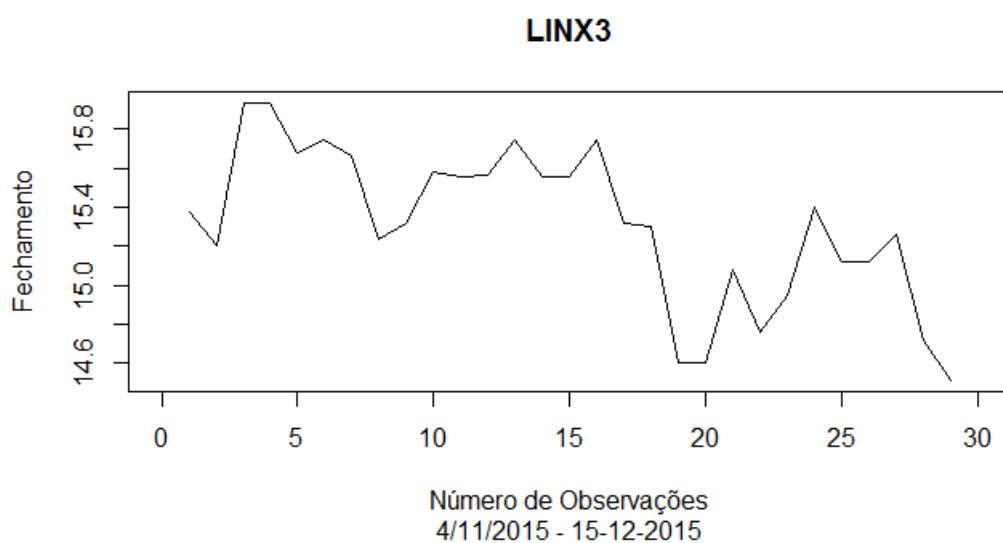


Figura 4.5.1: Gráfico Linx3

4.5.1 Identificação.

a) Plotagem da Série Original:

Como podemos observar na figura 4.5.1, esta não é uma série estacionária e por isso vamos diferenciá-la para que se torne estacionária.

b) Diferenciações:

Como podemos observar na figura 4.5.2, a série ainda não é estacionária, então vamos diferenciá-la novamente. Agora como podemos observar na figura 4.5.3 a

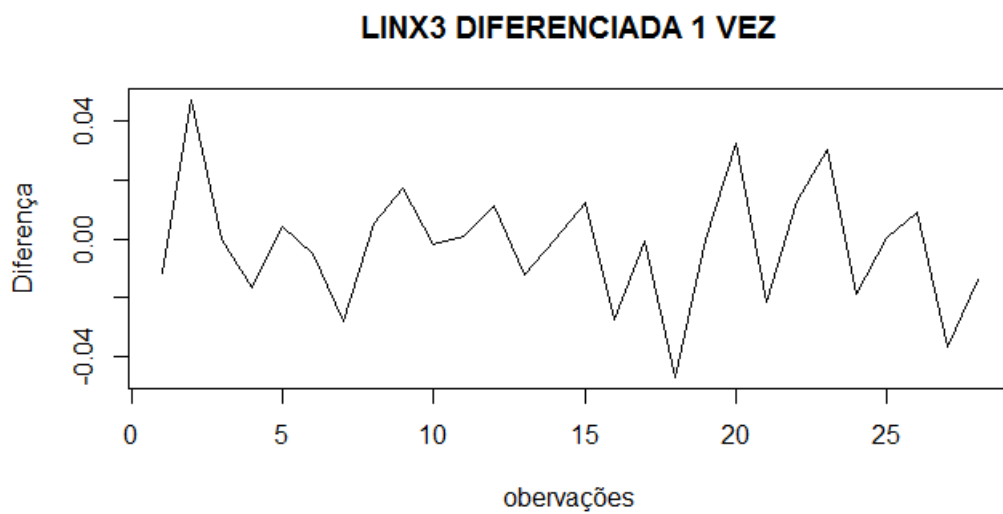


Figura 4.5.2: Gráfico Linx3 diferenciado 1 vez

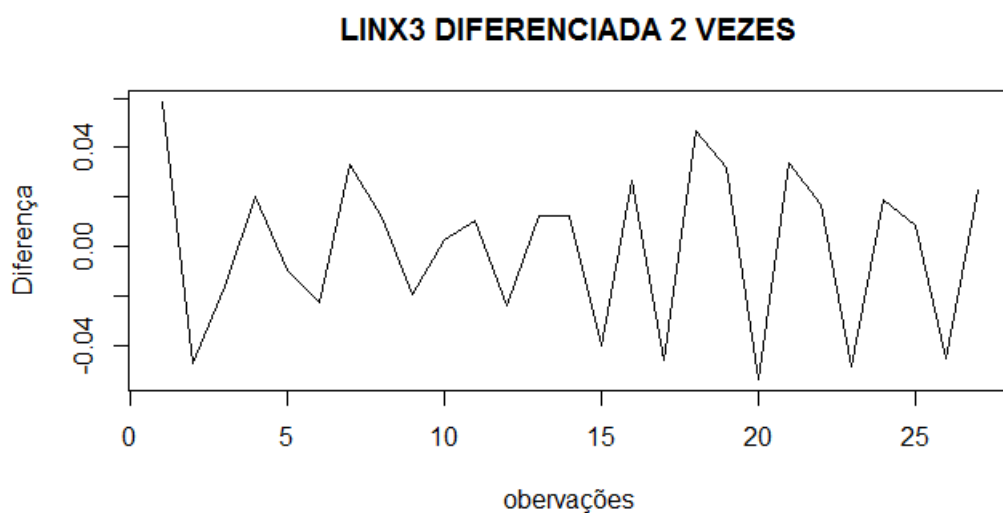


Figura 4.5.3: Gráfico Linx3 Diferenciado em lag 2

série está estacionária e portanto podemos continuar e encontrar os parâmetros p de AR e q de MA. Assim o parâmetro d de $AR(p)I(d)MA(q)$ é $d=2$.

c) Função de Autocorrelação:

Agora que a série está estacionária, basta aplicar os procedimentos para $AR(p)MA(q)$.

Como já vimos, a FAC serve para nos auxiliar no parâmetro (q) de MA que devemos utilizar na simulação. Veja as FAC da série original e diferenciada. Como

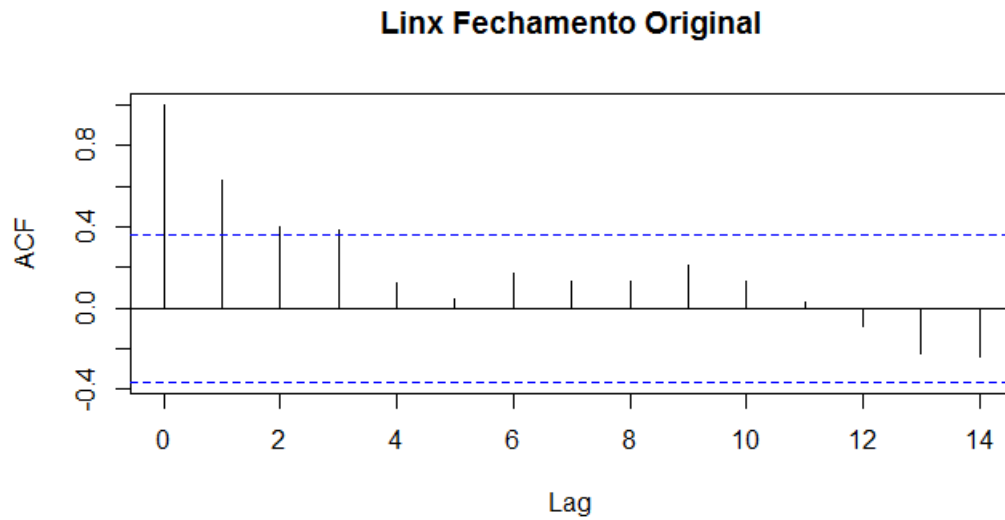


Figura 4.5.4: Autocorrelação da série original

podemos observar na figura 4.5.4, a série original sugere um parâmetro (q)=4. Porém, a série diferenciada (que nos interessa) sugere algo diferente, veja na figura 4.5.5. Portanto, para o nosso modelo, será utilizado $MA(q)=1$.

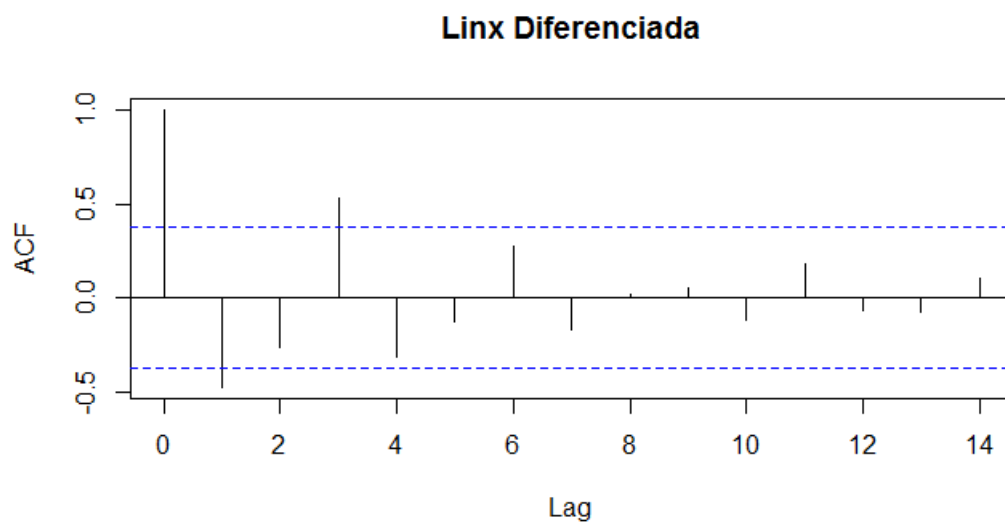


Figura 4.5.5: Autocorrelação da série diferenciada

d) Função de Autocorrelação Parcial:

Anteriormente vimos que as FACP nos ajudam a descobrir o parâmetro (p) de AR. Veja a seguir as facp da série original e diferenciada.

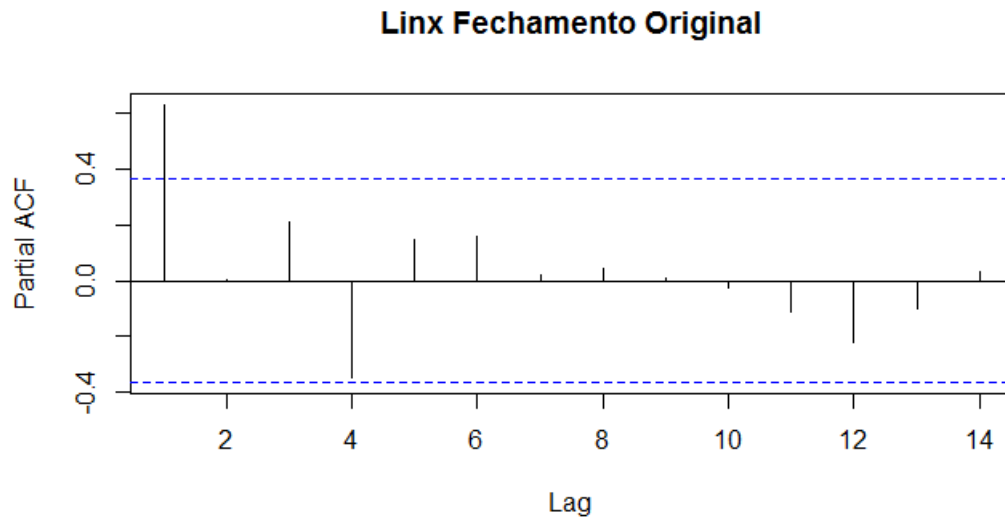


Figura 4.5.6: FACP da série original

A série original, na figura 4.5.6 apresenta uma autocorrelação parcial de 1, mas veremos que na série diferencia, onde de fato a facp funciona, é diferente de 1.

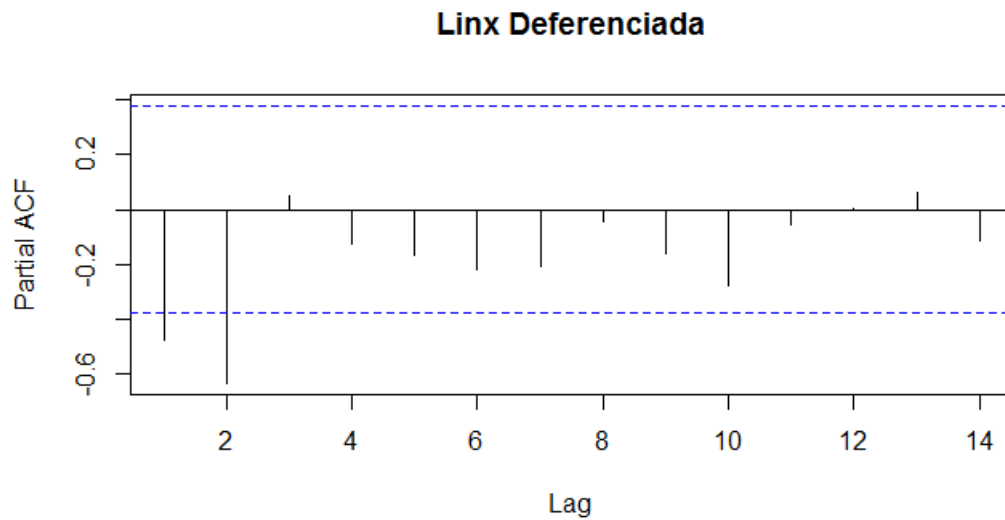


Figura 4.5.7: FACP da série diferenciada

Logo, pela figura 4.5.7, podemos observar que o parâmetro (p) sugerido pela facp da série diferenciada é 2. Com isso o nosso modelo está quase pronto e é do tipo:

$$X_t = \phi_1(X_{t-1}) + \phi_2(X_{t-2}) + \theta_1(X_{t-1}) + \varepsilon_t$$

4.5.2 *Estimação.*

e) Estimando os valores de ϕ_k e θ_k :

Como observamos na fac e na facp, os parâmetros (q) e (p) devem ser, respectivamente, 1 e 2. Assim podemos utilizar o Programa R para nos ajudar a estimar os valores dos coeficientes ϕ_k e θ_k . Para isso vamos digitar o seguinte comando.

```
arima(x = linxd, order = c(2, 0, 1))
```

onde `linxd` é a variável que guarda a série diferenciada. Ou ainda podemos aplicar diretamente na série original, onde os mesmos resultados apareceriam, mas deveríamos escrever da seguinte forma:

```
arima(x = linx$Adj.Close, order = c(2, 2, 1))
```

onde `linx$Adj.Close` é a variável que guarda a série original. Em ambos os casos o programa retorna a imagem 4.5.8:

```
Call:
arima(x = linxd, order = c(2, 0, 1))

Coefficients:
          ar1          ar2          ma1  intercept
      -0.9285  -0.7686   0.1735   -0.0011
s.e.   0.1818   0.1331   0.2654    0.0017
```

Figura 4.5.8: Coeficientes de ϕ_k e θ_k

4.5.3 Verificação.

f) Plotagem do modelo sobre a série original:

Em azul observamos a série original diferenciada e em vermelho o nosso modelo de série que tenta acompanhar o modelo original diferenciado.

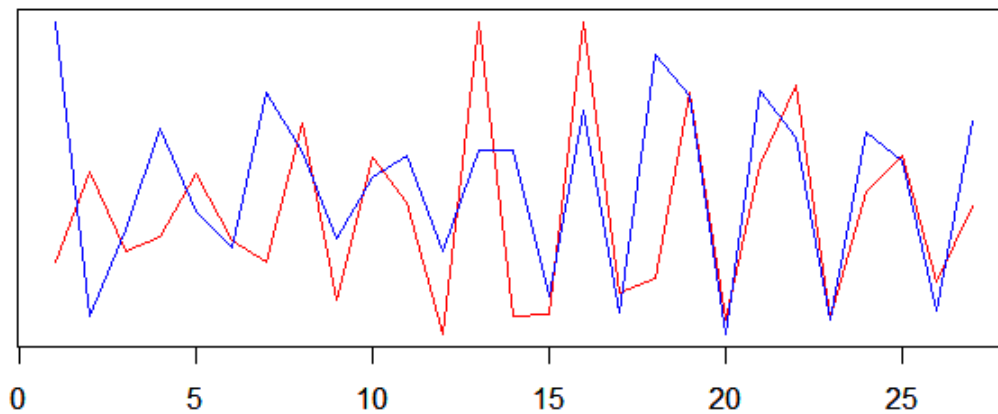


Figura 4.5.9: Simulação

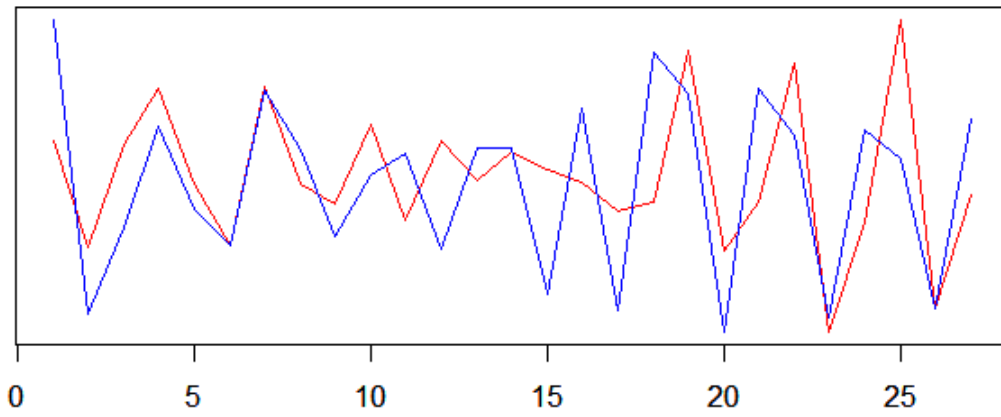


Figura 4.5.10: Simulação 2

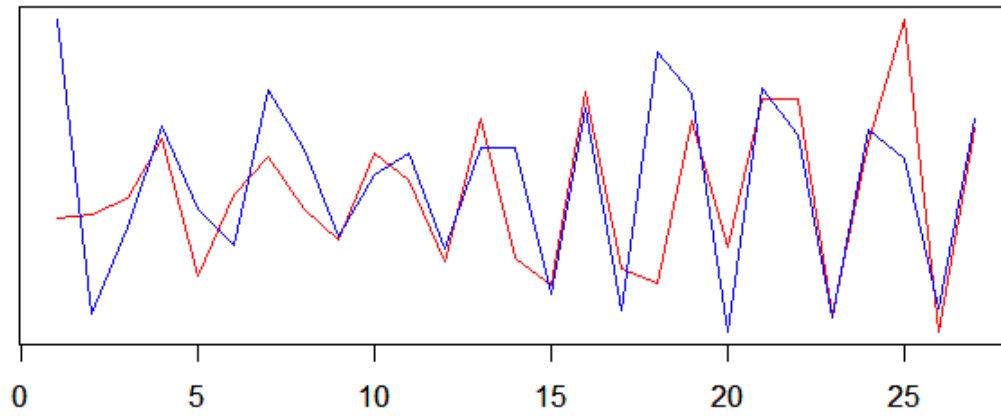


Figura 4.5.11: Simulação 3

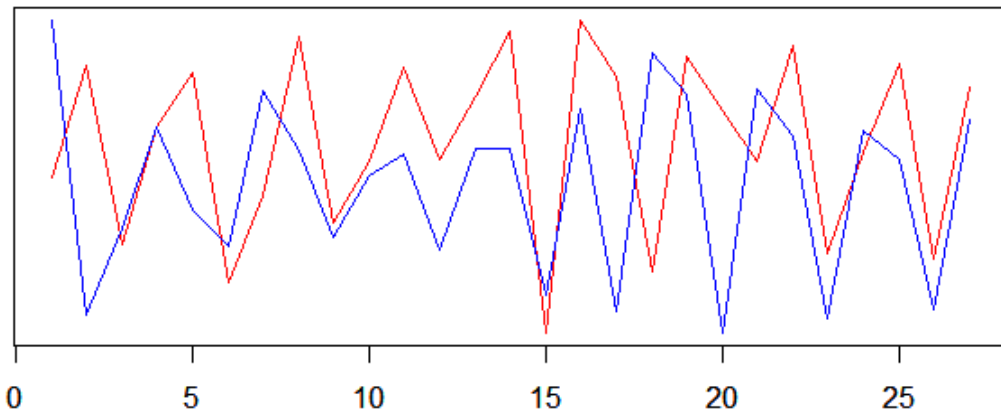


Figura 4.5.12: Simulação 4

A programação, as bibliotecas necessárias e o modo de instalação encontram-se no Apêndice C.

METODOLOGIA

A Escola do Estado apresenta uma série de dificuldades. Uma delas é o fato da maioria dos alunos carregarem uma defasagem ano após ano. Alguns alunos chegam ao 3º Ano do Ensino Médio sem saber realizar as 4 operações básicas. O plano de aula a seguir foi montado com o objetivo de reduzir a defasagem dos alunos e também com a intenção de promover a discussão sobre Bolsa de Valores e outros tipos de investimentos.

5.1 PLANO DE AULA

PLANO DE AULA DO ENSINO MÉDIO

1. IDENTIFICAÇÃO

Escola: Vila Magini II

Disciplina: Matemática

Carga horária: 45 h/a

Série: 3º ano

Ano: 2017

Professor: Roberto Sampaio Júnior

2. OBJETIVO GERAL A matemática comercial e financeira não é nova. Suas aplicações remontam de períodos anteriores a Cristo. A própria Bíblia Sagrada traz

referências de juros e de aplicações financeiras. O conceito de juros é antigo de acordo com os registros históricos. Essa conceituação apareceu quando o homem percebeu a relação entre o tempo e o dinheiro e seus reflexos na vida das pessoas e povos.

Os processos de acúmulo de capital e a desvalorização da moeda resultaram também na idéia dos juros. Registros antigos mostram que já existiam textos remotos que tratavam da distribuição de produtos agrícolas e de cálculos aritméticos baseados nessas transações.

Os Sumérios já utilizavam, adaptados à época, tipos de contratos legais, faturas, recibos, notas promissórias, crédito, juros e escrituras de venda. Portanto, espera-se que os alunos evoluam suas habilidades com frações, porcentagens, razão e proporção. Também é esperado que os alunos se motivem a guardar dinheiro e entender o poder dos juros compostos e sua importância histórica.

Esta aula se insere em blocos, são eles: 1 – Motivação; 2 – Preparo; 3 – Realização de cálculos; 4 – Revisão de resultados e socialização do conhecimento. As aulas pretendem organizar um ambiente provocador de aprendizagem e explorar as potencialidades pedagógicas da modelagem matemática. Nosso objetivo geral é abordar conteúdos como Razão e Proporção, Porcentagem, Juro, Taxa de Juros (simples e composto), Montante, Variação percentual, Juros compostos com taxa de juros variáveis. Através desses assuntos abordaremos o “Montante de uma sequência de depósitos”. Propomos uma discussão sobre aposentadoria como desencadeadora do processo, pois acreditamos que esse tipo de discussão possibilita momentos de reflexão entre os alunos, além do prospecto individual uma simulação de aposentadoria permite desenvolver. O objetivo final é a elaboração de uma fórmula simples para o cálculo de uma aposentadoria privada onde o montante a ser retirado (ou recebido mensalmente) e a data para esse acontecimento sejam variáveis escolhidas pelo indivíduo.

3. OBJETIVOS ESPECÍFICOS

Os objetivos em torno do PCN são:

- a) O estudo das funções permite ao aluno adquirir a linguagem algébrica como a linguagem das ciências, necessária para expressar a relação entre grandezas e **modelar** situações-problema, construindo modelos descritivos de fenômenos e permitindo várias conexões dentro e fora da própria matemática. Assim, a ênfase do estudo das diferentes funções deve estar no conceito

de função e em suas propriedades em relação às operações, na interpretação de seus gráficos como 3 e nas aplicações dessas funções. [1]

- b) Como o somatório dos juros compostos é a somatória de uma P.G., com relação às seqüências, é preciso garantir uma abordagem conectada à idéia de função, na qual as relações com diferentes funções possam ser analisadas. O estudo da progressão geométrica infinita com razão positiva e menor que 1 oferece talvez a única oportunidade de o aluno estender o conceito de soma para um número infinito de parcelas, ampliando sua compreensão sobre a adição e tendo a oportunidade de se defrontar com as idéias de convergência e de infinito (incluse para o dinheiro). Essas idéias foram e são essenciais para o desenvolvimento da ciência, especialmente porque permitem explorar regularidades e padrões.
- c) A Estatística e a Probabilidade devem ser vistas, então, como um conjunto de idéias e procedimentos que permitem aplicar a Matemática em questões do mundo real, mais especialmente aquelas provenientes de outras áreas. Devem ser vistas também como formas de a Matemática quantificar e interpretar conjuntos de dados ou informações que não podem ser quantificados direta ou exatamente. Cabe à Estatística, por exemplo, analisar a intenção de voto em uma eleição ou o possível êxito do lançamento de um produto no mercado, antes da eleição em si e da fabricação do produto. Isso é feito através da pesquisa estatística, que envolve amostras, levantamento de dados e análise das informações obtidas. Da mesma forma, a Probabilidade acena com resultados possíveis, mas não exatos. Ao afirmar que o resultado 1 tem $1/6$ de probabilidade no lançamento de um dado, não há certeza de que em seis lançamentos do dado o número 1 sairá exatamente uma vez. Assim como ao afirmarmos que determinado tratamento médico tem 90% de probabilidade de cura para uma doença, não garante que em um grupo de 10 pessoas submetidas a este tratamento exatamente uma pessoa continuará doente. [1]

4. CONTEÚDO PROGRAMÁTICO Duração das atividades

1 Bimestre – cada aula com 50 minutos.

Conhecimentos prévios trabalhados pelo professor com o aluno. Para esta aula, é importante que os alunos entendam multiplicação e divisão de números racio-

nais, utilização de tabelas para organizar dados, conhecimentos sobre variáveis dependentes e independentes.

Atividades

Em um primeiro momento sugerimos uma discussão sobre a atual forma de aposentadoria proposta pelo governo. Acreditamos que essa discussão motivará os alunos para a atividade experimental que vem a seguir. Após esta discussão, sugerimos a idealização de uma aposentadoria e a partir da dificuldade de uma simulação da sistematização de dados abordaremos os conteúdos a serem atingidos. Para finalizar, em um terceiro momento sugerimos a discussão dos resultados do experimento com a turma e, a partir deles, uma abordagem do conceito de progressão geométrica.

Aula 1 – O Brasil e sua aposentadoria. Nesta aula abordaremos as regras da aposentadoria no Brasil, os valores oferecidos por ela e o quanto de cada pagamento que temos de contribuir. Com algumas contas básicas de porcentagem e soma, a intenção é que os alunos percebam as desvantagens de esperar pela aposentadoria e assim levantar alguma estratégia para a obtenção da autonomia financeira. Foi observado, durante as aplicações, que a poupança seria o meio mais utilizado para obtenção dessa autonomia. Recursos necessários: Internet para a pesquisa ou o professor pode levar impressos.

Aula 2 e 3 – A poupança e a SELIC. Nesta aula abordaremos as regras da poupança no Brasil, os valores oferecidos por ela e o quanto de rendimento ela nos traz. Pela pesquisa, os alunos observarão que há regras para a poupança e que conforme a SELIC ela muda. O professor deve pedir então aos alunos que pesquisem o que é a SELIC e o que ela tem a ver com a inflação, surgindo o novo questionamento. O que é inflação? O que leva a um raciocínio mais crítico sobre a poupança, nos empurrando a uma nova pesquisa: Fontes de rendimento além da poupança. *Aula 4 e 5 – Fontes de rendimento além da Poupança.* No Brasil, existem diversos títulos comercializados no mercado financeiro que possuem característica de renda fixa. Dentre os que são oferecidos para pessoas físicas, os mais comuns são CDB, RDB, LC, LCI, LCA, CRA, CRI, FIDC, LF, DPGE, Debêntures e Tesouro Direto. Após essa pesquisa o professor deve perguntar como calcular a diferença de rendimento entre eles (CDB, RDB, LCI e LCA) e a poupança após 1 ano. Isso leva a uma sequência de multiplicações muito extensa onde o professor deve questionar se não há outro método, mais fácil, para a obtenção desses resultados. Assim se inicia a parte teórica e uma sequência

de exercícios proposta no livro Fundamentos da Matemática Elementar Vol. 11 – Mat. Comercial/Financeira e Estatística Descritiva. Iremos avançar por toda a parte de porcentagem ao passo que a fórmula dos juros compostos, gerado por uma P.G, é demonstrada pelo professor e em seguida é apresentado o site (www.calculadorabc) para que possam explorar a fórmula.

Aula 6, 7 e 8 – Exercícios sobre Razão e Proporção. Como treinamento e aperfeiçoamento das técnicas matemáticas uma sequência de exercícios foi proposta. As aulas 6, 7 e 8 foram utilizadas para a realização dos exercícios sobre razão e proporção. Ex: 1 – 18 do livro de Fundamentos.

Aula 9, 10, 11 e 12 – Exercícios sobre GDP e GIP. Como treinamento e aperfeiçoamento das técnicas matemáticas uma sequência de exercícios foi proposta. As aulas 9, 10, 11 e 12 foram utilizadas para a realização dos exercícios sobre Grandezas Diretamente Proporcionais e Grandezas Inversamente Proporcionais. Ex: 19 – 35 do livro de Fundamentos. Daremos atenção especial ao Exercício 35 onde há uma profunda manipulação das variáveis. Este exercício dá início ao nosso objetivo final. Durante a correção desse exercício o professor tem a oportunidade de mostrar o nível de sutileza que esperamos que os alunos desenvolvam até o final do projeto.

Aula 13, 14, 15 e 16 – Exercícios sobre porcentagem. Como treinamento e aperfeiçoamento das técnicas matemáticas uma sequência de exercícios foi proposta. As aulas 13, 14, 15 e 16 foram utilizadas para a realização dos exercícios sobre Porcentagens. Ex: 36 – 55 do livro de Fundamentos. Exercícios de pequena dificuldade

Aula 17, 18, 19 e 20 – Exercícios sobre porcentagem. Como treinamento e aperfeiçoamento das técnicas matemáticas uma sequência de exercícios foi proposta. As aulas 17, 18, 19 e 20 foram utilizadas para a realização dos exercícios sobre Porcentagens. Ex: 56 – 75 do livro de Fundamentos. Exercícios de média dificuldade.

Aula 21 e 22 – Exercícios sobre porcentagem. Como treinamento e aperfeiçoamento das técnicas matemáticas uma sequência de exercícios foi proposta. As aulas 21 e 22 foram utilizadas para a realização dos exercícios sobre Porcentagens. Ex: 76 – 84 do livro de Fundamentos. Exercícios de grande dificuldade.

Aula 23 – Discussão sobre Inflação e Acúmulo de Capital. Após a realização de tantos exercícios era esperado um mínimo domínio sobre manipulação de variáveis,

para que haja o entendimento da demonstração da fórmula dos juros compostos e da demonstração da fórmula do montante de capital após n depósitos recorrentes a uma taxa fixa i . A necessidade da demonstração surge quando necessitamos calcular o tempo que demoraria à alcançar 1 milhão de reais e quanto deveríamos depositar por mês.

Aula 24, 25, 26 e 27 – Exercícios sobre Inflação. Como treinamento e aperfeiçoamento das técnicas matemáticas uma sequência de exercícios foi proposta. As aulas 24, 25, 26 e 27 foram utilizadas para a realização dos exercícios sobre Inflação. Ex: 111 – 127 do livro de Fundamentos. Exercícios de média dificuldade.

Aula 28 e 29 – Exercícios sobre Regimes de Capitalização. Como treinamento e aperfeiçoamento das técnicas matemáticas uma sequência de exercícios foi proposta. As aulas 28 e 29 foram utilizadas para a realização dos exercícios sobre R.C.. Ex: 143 – 149 do livro de Fundamentos. Exercícios de média dificuldade.

Aula 30 – Exercícios sobre Juros Simples. Como treinamento e aperfeiçoamento das técnicas matemáticas uma sequência de exercícios foi proposta. A aula 30 foi utilizada para a realização dos exercícios sobre Juros Simples e demonstração da fórmula. Ex: 156 – 161 do livro de Fundamentos. Exercícios de média dificuldade.

Aula 31, 32, 33 e 34 – Exercícios sobre Juros Compostos. Como treinamento e aperfeiçoamento das técnicas matemáticas uma sequência de exercícios foi proposta. As aulas 31, 32, 33 e 34 foram utilizadas para a realização dos exercícios sobre Juros Compostos e demonstração da fórmula. Ex: 188 – 207 do livro de Fundamentos. Exercícios de média dificuldade.

Aula 35 e 36 – Exercícios sobre o Valor Atual de um Conjunto de Capitais. Como treinamento e aperfeiçoamento das técnicas matemáticas uma sequência de exercícios foi proposta. As aulas 35 e 36 foram utilizadas para a realização dos exercícios sobre Valor Atual de um Conjunto de Capitais. Ex: 226 - 234 do livro de Fundamentos. Exercícios de média dificuldade.

Aula 37, 38 e 39 – Exercícios sobre Sequência uniforme de Pagamentos. Como treinamento e aperfeiçoamento das técnicas matemáticas uma sequência de exercícios foi proposta. As aulas 37, 38 e 29 foram utilizadas para a realização dos exercícios sobre Sequência uniforme de Pagamentos e demonstração da fórmula. Ex: 235 – 246 do livro de Fundamentos. Exercícios de média dificuldade.

Aula 40, 41 e 42 – Exercícios sobre Montante de uma Sequência uniforme de depósitos. Como treinamento e aperfeiçoamento das técnicas matemáticas uma sequência de exercícios foi proposta. As aulas 40, 41 e 42 foram utilizadas para a realização dos exercícios sobre Montante de uma sequência uniforme de depósitos e demonstração da fórmula. Ex: 247 – 254 do livro de Fundamentos. Exercícios de média dificuldade.

Aula 43 – Exploração no site do Banco Central. Ao encerrar a lista de exercícios e fazer as demonstrações das fórmulas, nosso objetivo foi concluído. Agora fica bastante interessante mostrar o “site” do Banco Central (<https://www3.bcb.gov.br/CALCIDADA0/jsp/index.jsp>) afim de fazer simulações e mostrar na aba “metodologia” que o BC usa as mesmas fórmulas que desenvolvemos.

4. Estratégias e recursos da aula

Estratégias:

- Perguntas guiadas pelo professor, com o intenção de promover a habilidade de manipulação de variáveis.
- Lista de exercícios de Matemática Comercial.
- Lista de exercícios de Matemática Financeira.
- Tentativas de Modelagem.

Recursos:

- Livro: Fundamentos da Matemática Elementar Vol. 11 – Mat. Comercial/Financeira e Estatística Descritiva. [10]
- Calculadora.

5.1.1 Aproveitamento dos alunos.

Após a aplicação do Plano de aula observou-se que os objetivos foram parcialmente atingidos. A grande maioria dos alunos não recuperou a desafagem em relação aos outros anos, porém houve muito interesse do todos sobre o assunto de investimentos e os caminhos para uma maior liberdade financeira. A figura 5.1.1 mostra um esboço de

uma aluna ao tentar demonstrar a fórmula de depósitos recorrentes. As figuras 5.1.2 e 5.1.3.

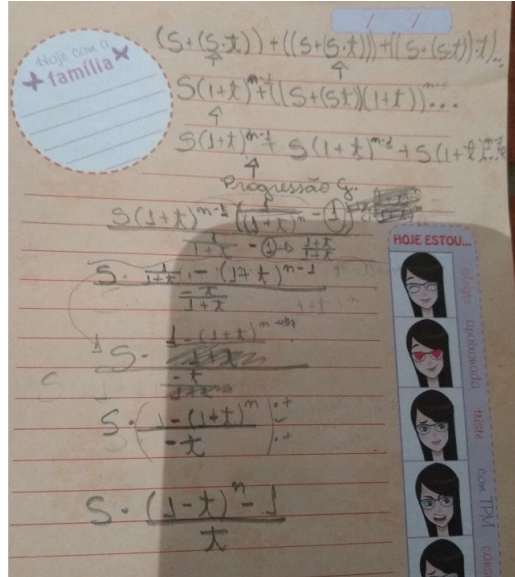


Figura 5.1.1: Esboço de uma aluna ao tentar demonstrar a fórmula de depósitos recorrentes.

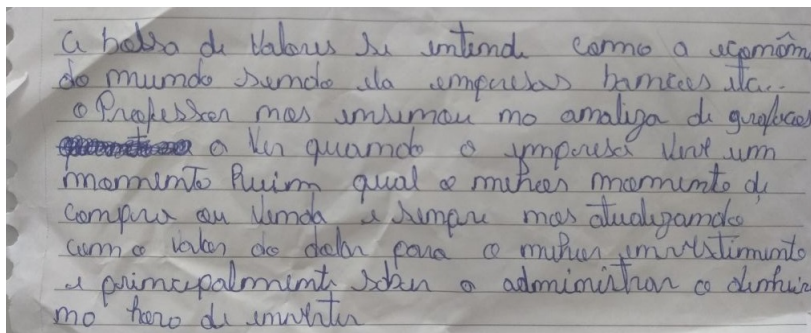


Figura 5.1.2: Depoimento do Aluno 1

5.2 YAHOO FINANCE.

Nesta seção vamos mostrar como extrair os dados de uma empresa a partir do “site” do “Yahoo Finance” e após a extração vamos mostrar como manipular os dados para realizar uma modelagem ARIMA no programa R Studio.

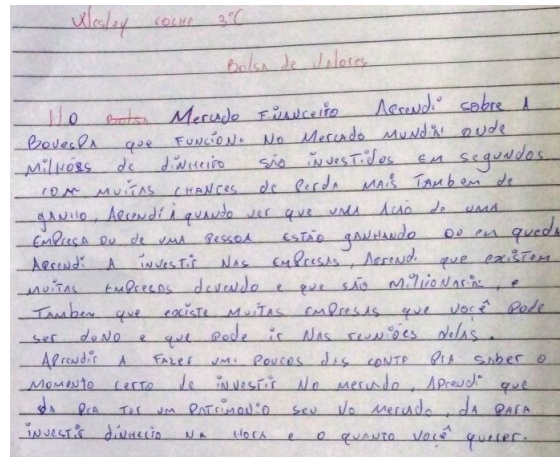


Figura 5.1.3: Depoimento do Aluno 2

5.2.1 Abrindo o site.

Primeiramente vamos abrir nosso navegador e digitar

<https://br.financas.yahoo.com>



Figura 5.2.1: Site do Yahoo Finance.

5.2.2 Escolhendo a Ação.

No campo indicado na figura 5.2.2 é onde encontramos a ação desejada. Vamos digitar “PETR3” e assim vamos encontrar as ações da Petrobrás.



Figura 5.2.2: Campo de digitação.

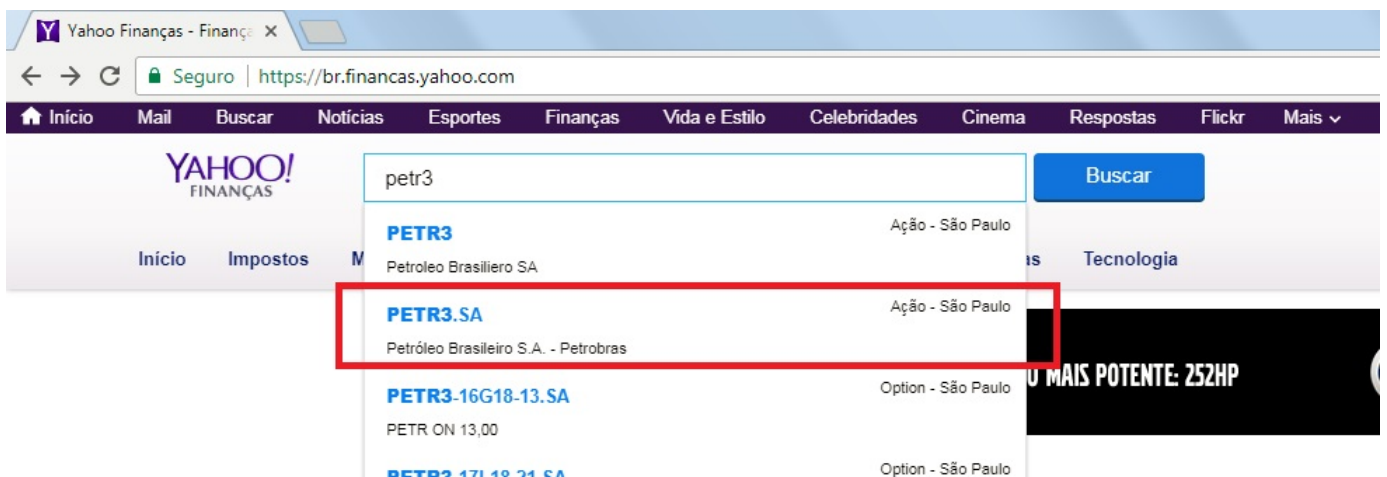


Figura 5.2.3: Escolhendo a Ação.

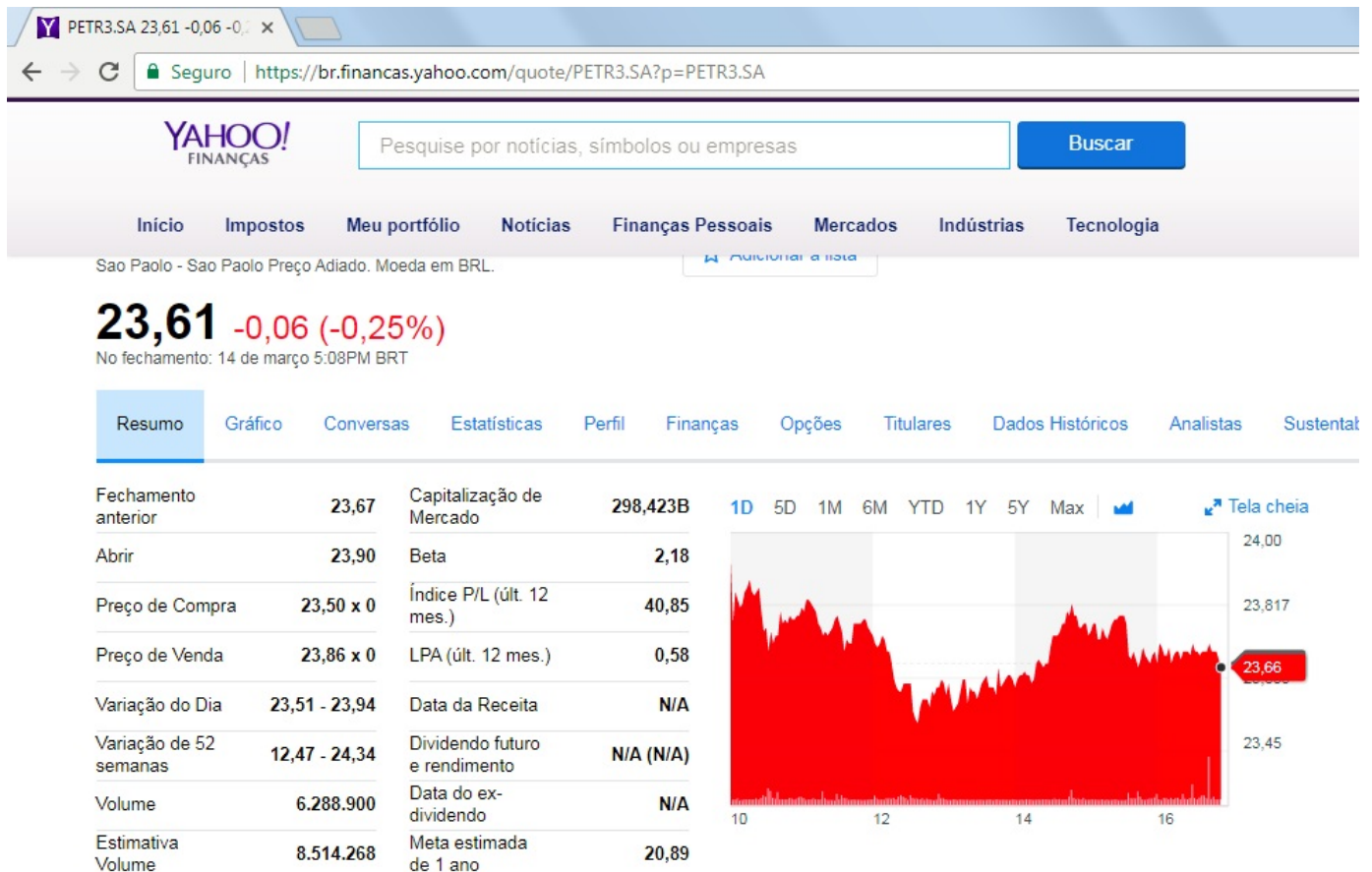


Figura 5.2.4: Escolhendo a Ação.

5.2.3 Indo para os Dados Históricos.

Após a escolha da ação o navegador abrirá uma tela como a da figura 5.2.4. E devemos clicar em Dados Históricos, como indicado na figura 5.2.5

5.2.4 Extraindo os Dados Históricos.

Agora estaremos em uma página como a da figura 5.2.6. O próximo passo é alterar o período. Para nosso exemplo vamos alterar o período de início para 12/15/2017 e o de término para 22/01/2018. Logo em seguida clicamos em concluído e depois

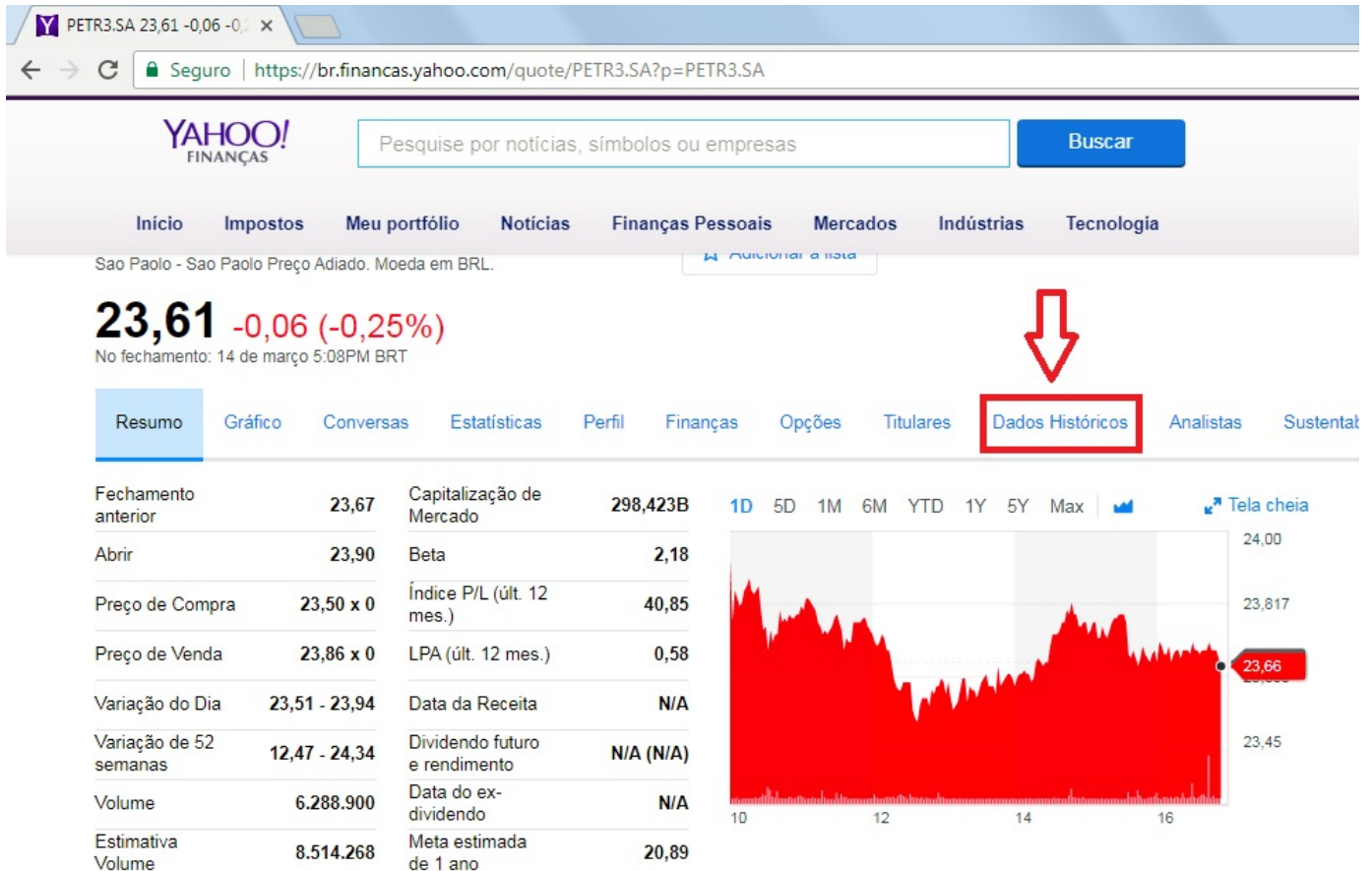


Figura 5.2.5: Onde clicar.

clicamos em aplicar. Como na figura 5.2.7 Após clicar em aplicar o “site” ira mostrar uma série de valores e para fazer o “download” basta clicar em “Fazer download dos dados”. Como na figura 5.2.8.

Agora temos um arquivo do tipo CSV em Excel. Para manipular os dados vamos usar o R Studio.

5.3 R STUDIO

Vamos mostrar agora o passo a passo para a modelagem ARIMA através do R Studio. Antes disso, vamos entender alguns detalhes do arquivo CSV que acabamos de baixar.

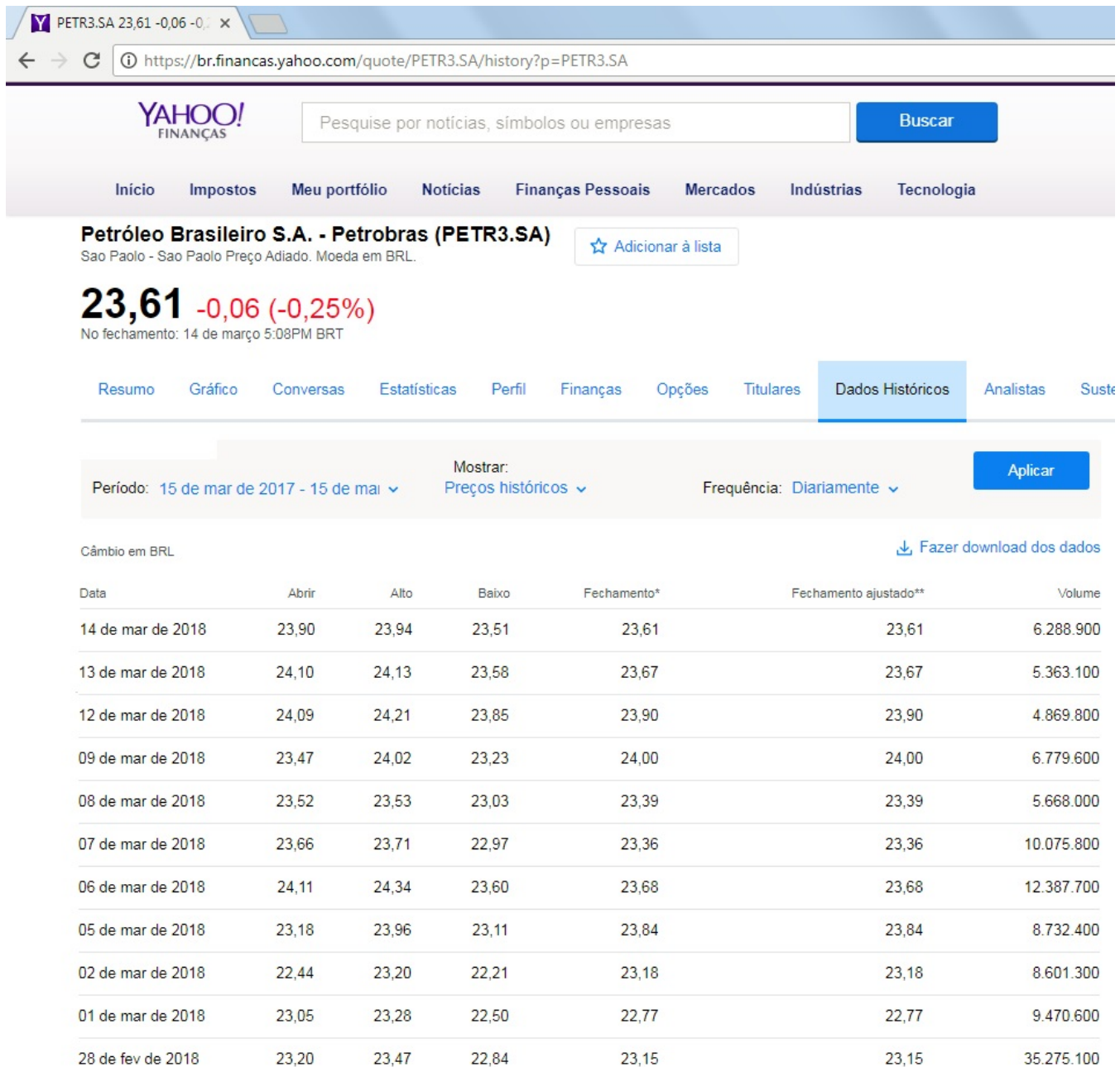


Figura 5.2.6: Os Dados Históricos.

5.3.1 O Arquivo CSV.

CSV (Comma-Separated Values) é um formato simples de armazenamento, que agrupa as informações de arquivos de texto em planilhas, para as trocas de dados

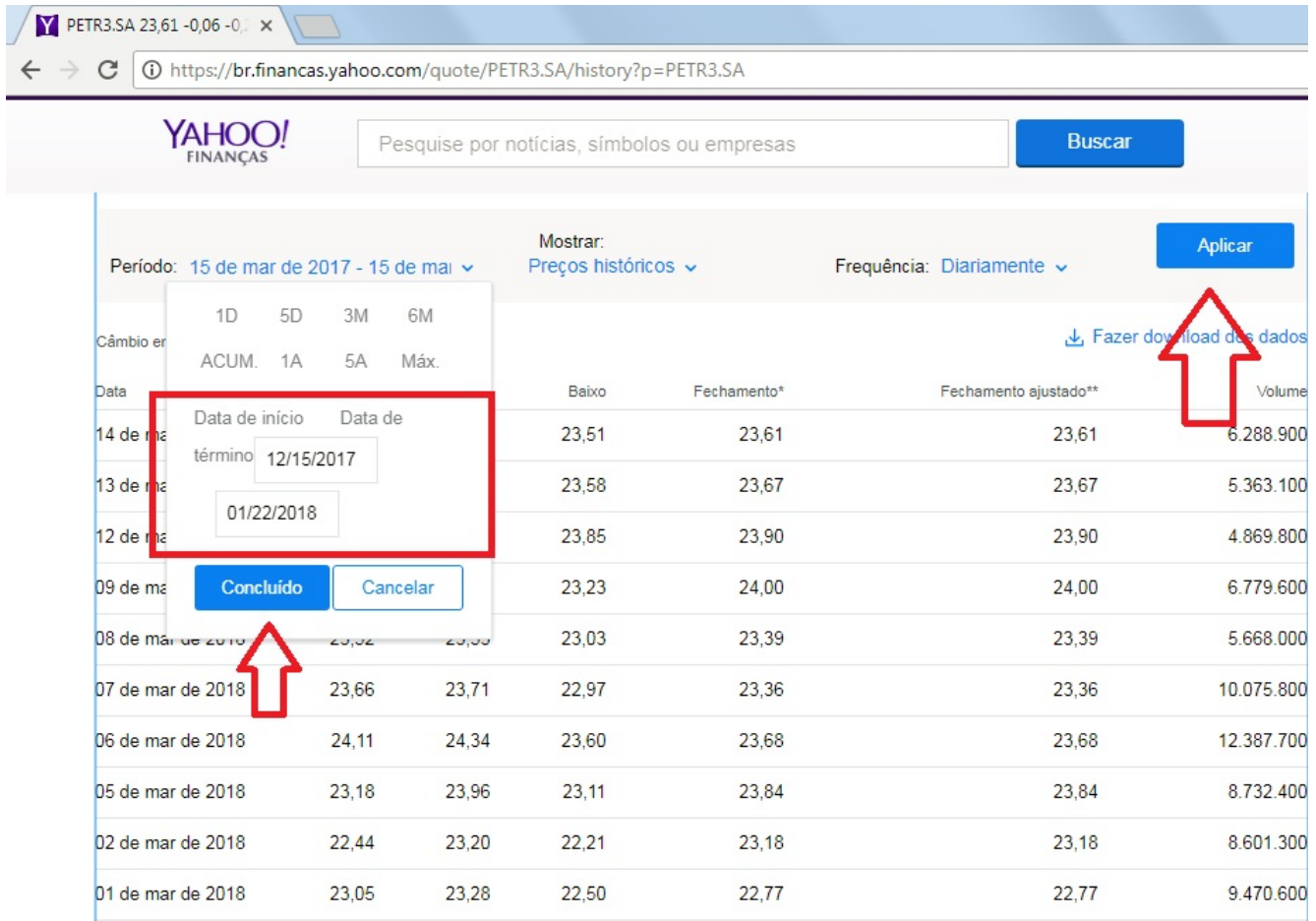


Figura 5.2.7: Alterando o período.

com um banco de dados ou uma planilha entre aplicativos. Cada linha em um texto CSV representa uma linha em uma planilha. Cada célula é geralmente separada por vírgula ou um outro caractere, como tabulador. Veja na figura 5.3.1.

O arquivo mostra muitas informações. As informações que vão nos interessar são os preços de fechamento. No arquivo essas informações são os penúltimos números, na coluna “Adj Close”.

Período: 15 de dez de 2017 - 22 de jan ▾ Mostrar: Preços históricos ▾ Frequência: Diariamente ▾ Aplicar

Câmbio em BRL

Data	Abrir	Alto	Baixo	Fechamento*	Fechamento ajustado**	Volume
22 de jan de 2018	19,21	19,57	19,21	19,51	19,51	7.741.500
19 de jan de 2018	19,50	19,62	19,22	19,31	19,31	4.134.700
18 de jan de 2018	19,47	19,66	19,13	19,44	19,44	8.160.700
17 de jan de 2018	18,92	19,47	18,86	19,47	19,47	6.517.400
16 de jan de 2018	18,45	18,93	18,43	18,79	18,79	7.189.900
15 de jan de 2018	18,31	18,55	18,31	18,46	18,46	2.308.600
12 de jan de 2018	18,10	18,41	18,02	18,36	18,36	5.516.100
11 de jan de 2018	17,89	18,27	17,81	18,22	18,22	10.075.000
10 de jan de 2018	18,00	18,11	17,76	17,78	17,78	4.720.000
09 de jan de 2018	18,06	18,27	18,01	18,12	18,12	6.570.100
08 de jan de 2018	17,61	18,12	17,61	18,12	18,12	4.763.400

PETR3.SA.csv

Figura 5.2.8: Download do período.

5.3.2 Abrindo as bibliotecas do R Studio.

Ao abrir o R Studio iremos nos deparar com uma tela como a da figura 5.3.2.

A tela de comando é a tela destacada na figura 5.3.3.

Para abrir as bibliotecas devemos digitar os comandos que estão na figura 5.3.4.

Para abrir o arquivo CSV necessário devemos digitar o comando em destaque na figura 5.3.5. Para isso devemos saber o local onde o arquivo foi salvo. Para evitar

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Date,Open,High,Low,Close,Adj Close,Volume															
2	2017-12-15,15.810000,16.059999,15.560000,15.560000,15.560000,7398800															
3	2017-12-18,15.780000,16.030001,15.730000,15.840000,15.840000,2158900															
4	2017-12-19,15.820000,15.850000,15.630000,15.760000,15.760000,028400															
5	2017-12-20,15.800000,15.000000,15.770000,15.960000,15.960000,848600															
6	2017-12-21,16.000000,16.650000,15.990000,16.650000,16.650000,998700															

Figura 5.3.1: Tabela CSV.

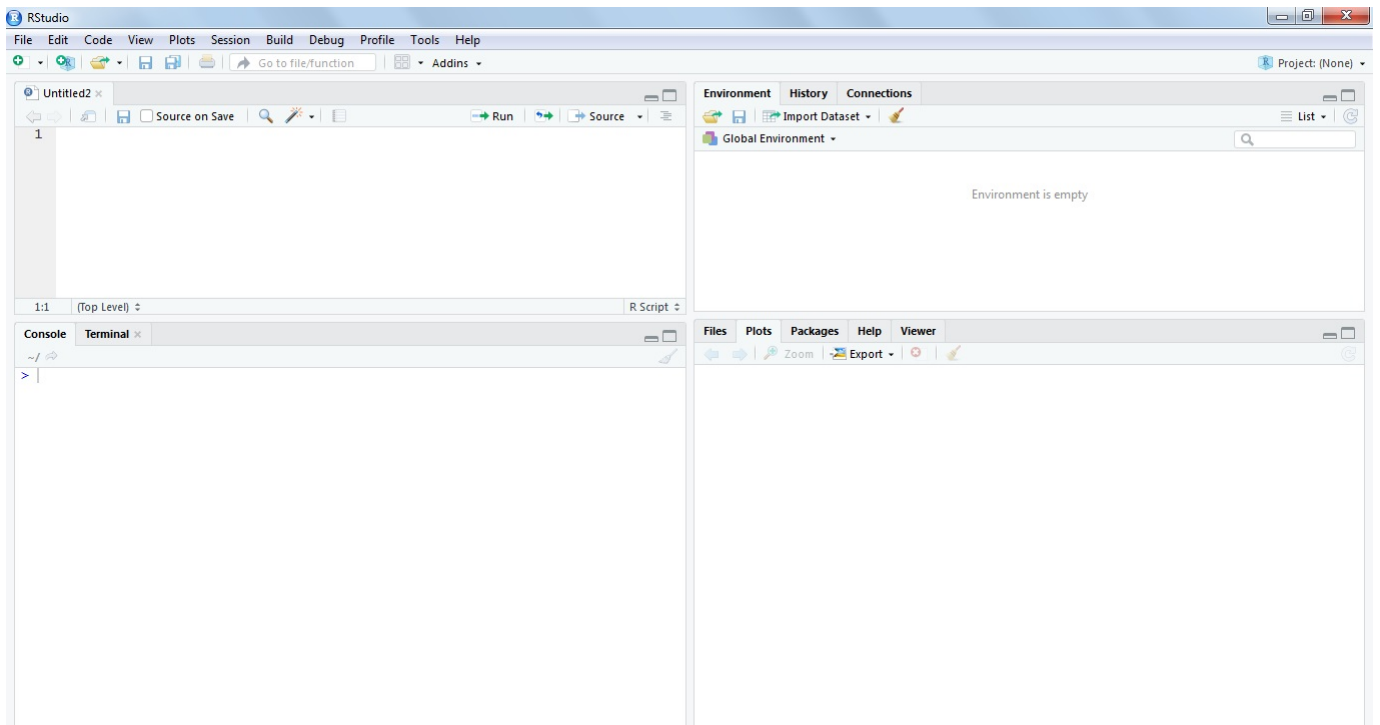


Figura 5.3.2: R Studio - Tela inicial.

erros, não salve o arquivo com pontos ou traços. Em nosso exemplo vamos renomear o arquivo CSV para “petr3”. Então, no caso deste exemplo, vamos digitar

```
petr = read.csv("C : /Users/User/Desktop/Tcc/PythonFiles/petr3.csv")
```

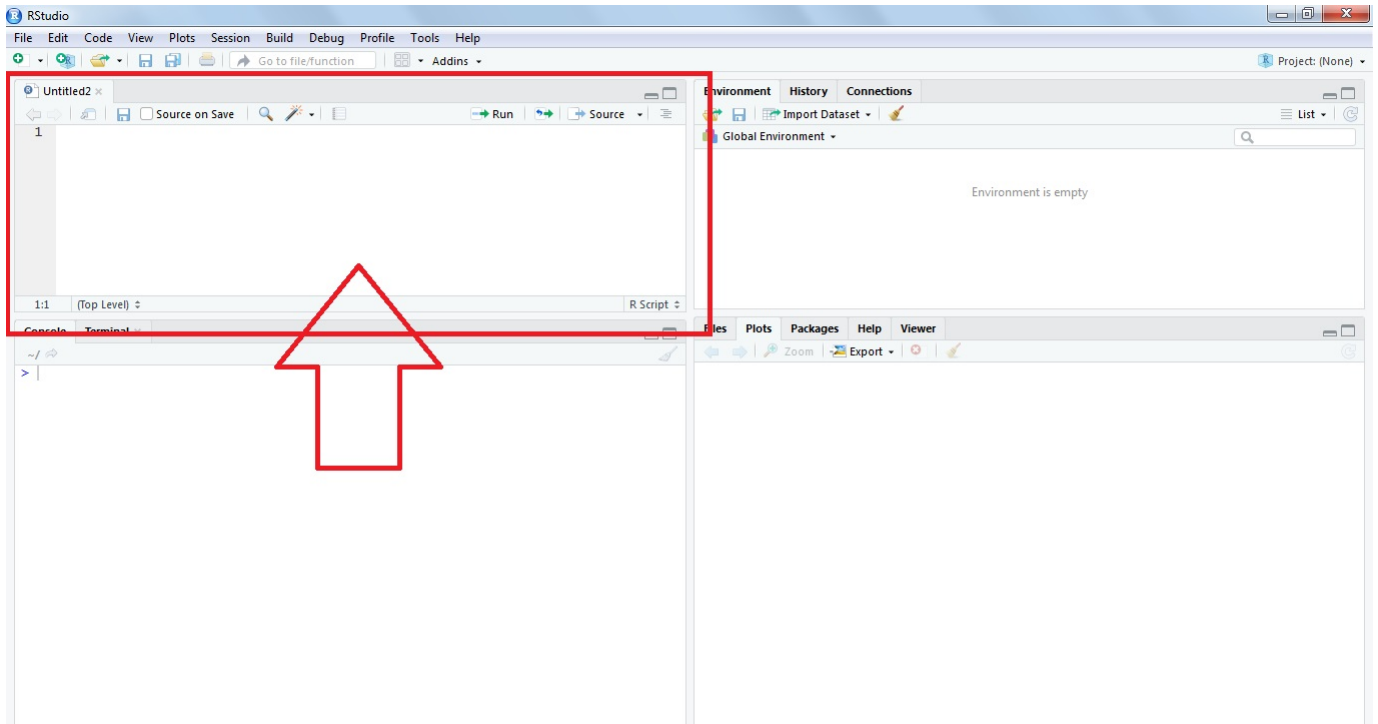


Figura 5.3.3: R Studio - Tela de comando.

5.3.3 *Eliminando Alguns Dados*

O arquivo CSV que baixamos da internet contém muitas informações. Algumas dessas informações não vamos utilizar, para isso vamos usar o comando NULL para

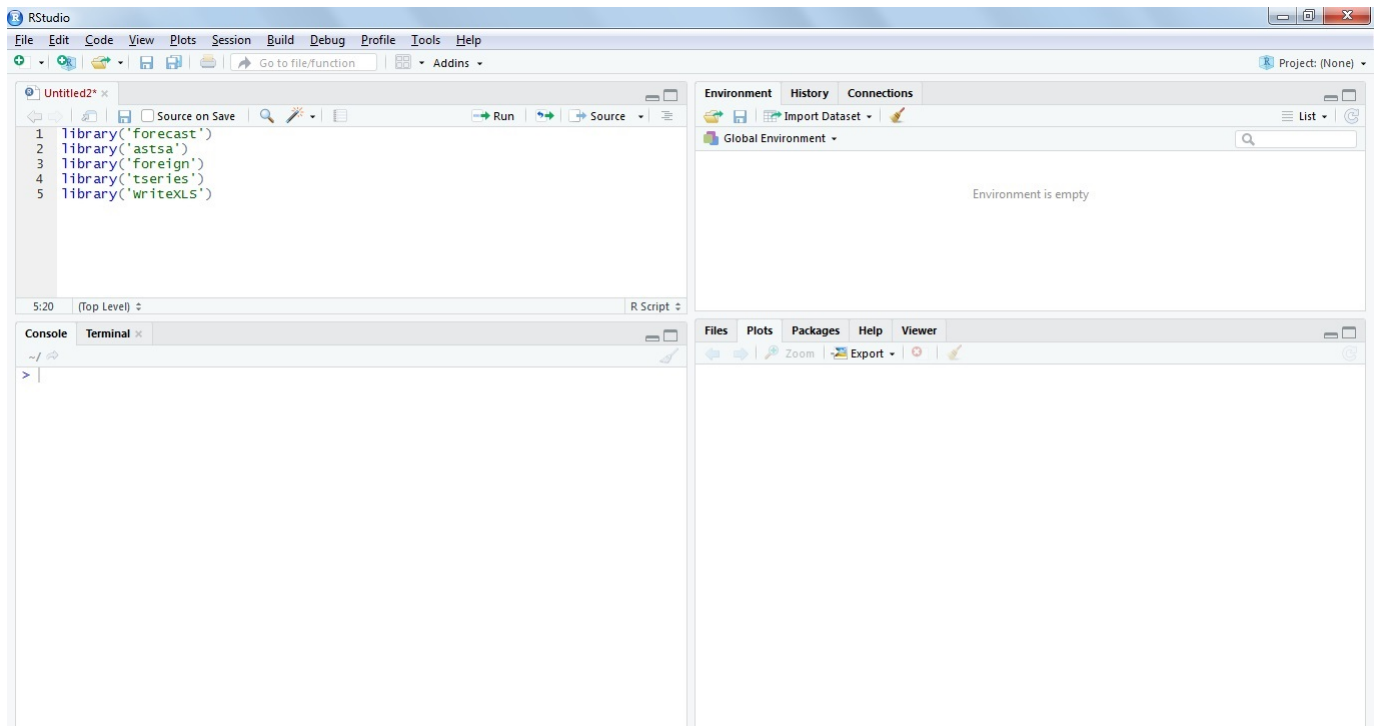


Figura 5.3.4: R Studio - Abrindo as bibliotecas.

apagar da planilha os dados desnecessários aos nossos estudos. Acompanhe na figura 5.3.6

Para entender o que está acontecendo vamos apertar o botão “Re-run” em destaque na figura 5.3.7.

Como podemos observar, uma tabela foi criada. Esta tabela contém somente os valores de fechamento das ações PETR3. Veja na figura 5.3.8.

5.3.4 *Plotando o gráfico original.*

Para plotar o grafico original, devemos digitar o seguinte comando:

```
plot(linx$Adj.Close, type = 'l', xlim = c(0, 26), main = 'PETR3')
```

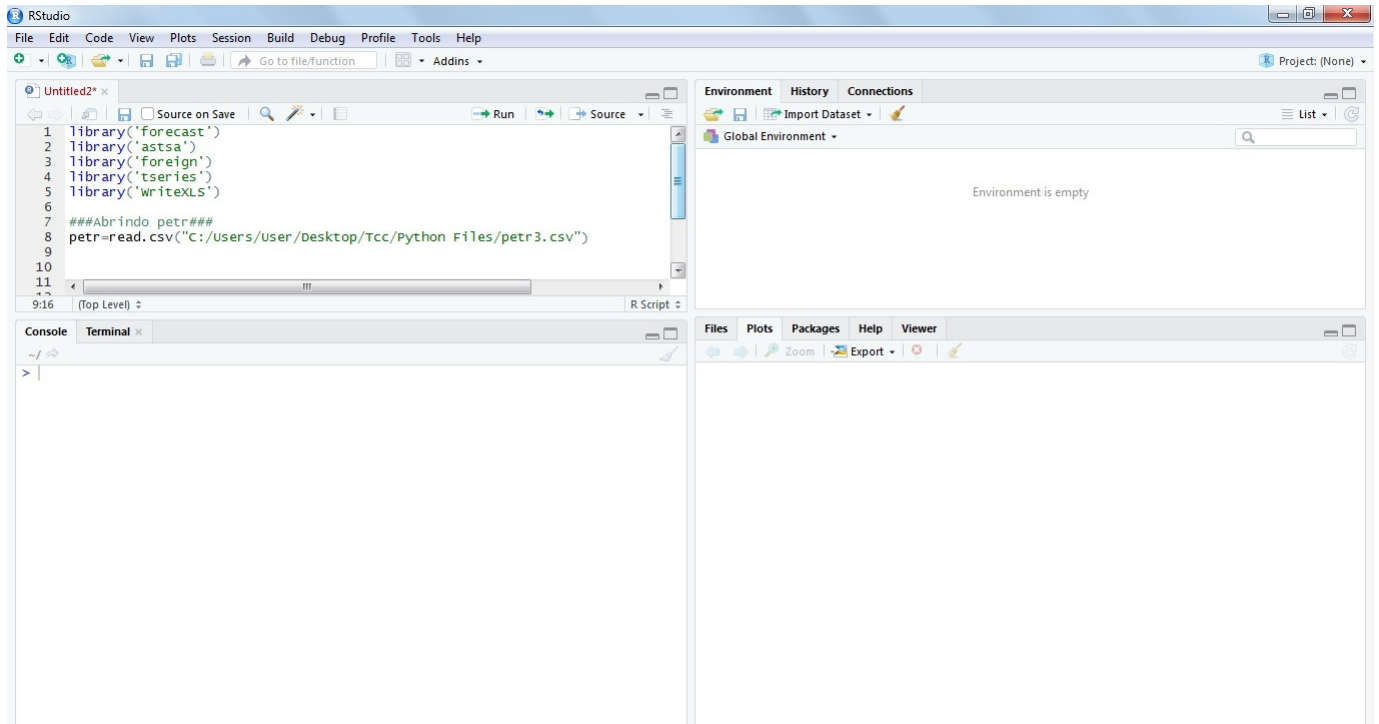



Figura 5.3.5: R Studio - Abrindo o Arquivo CSV.

O subcomando “type” dentro do comando plot faz a função de traçar o gráfico com linhas. Agora vamos clicar no botão “Re-run” para o gráfico ser plotado. Veja nas figuras 5.3.9 e 5.3.10.

5.3.5 Diferenciando a série.

Como podemos observar, o gráfico plotado não é estacionário. Para isso vamos utilizar o comando:

$$\text{diff}(x = \log(\text{linx}\$Adj.Close), \text{differences} = 1)$$

Para diferenciar uma vez e então vamos plotar o gráfico. Veja as figuras 5.3.11 e 5.3.12

Como agora julgamos a série como estacionária, vamos aos próximos passos.

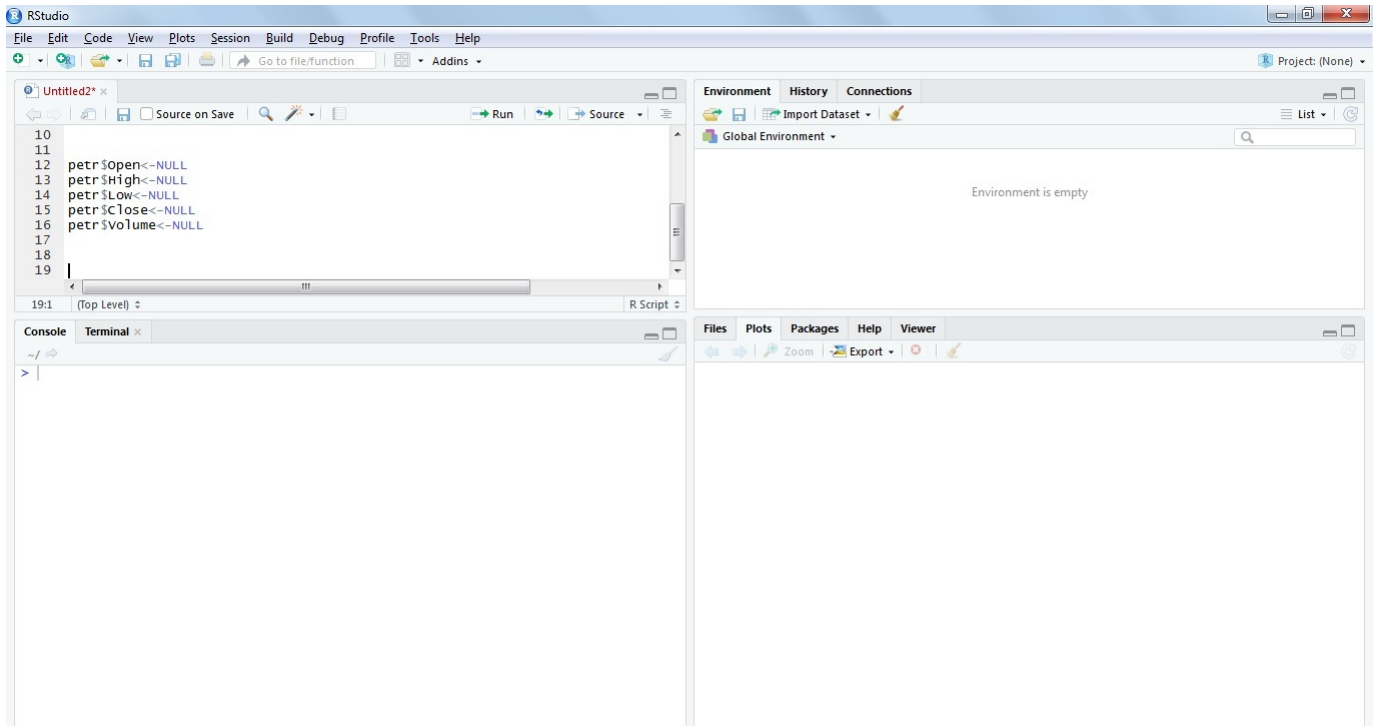


Figura 5.3.6: R Studio - Eliminando dados do Arquivo CSV.

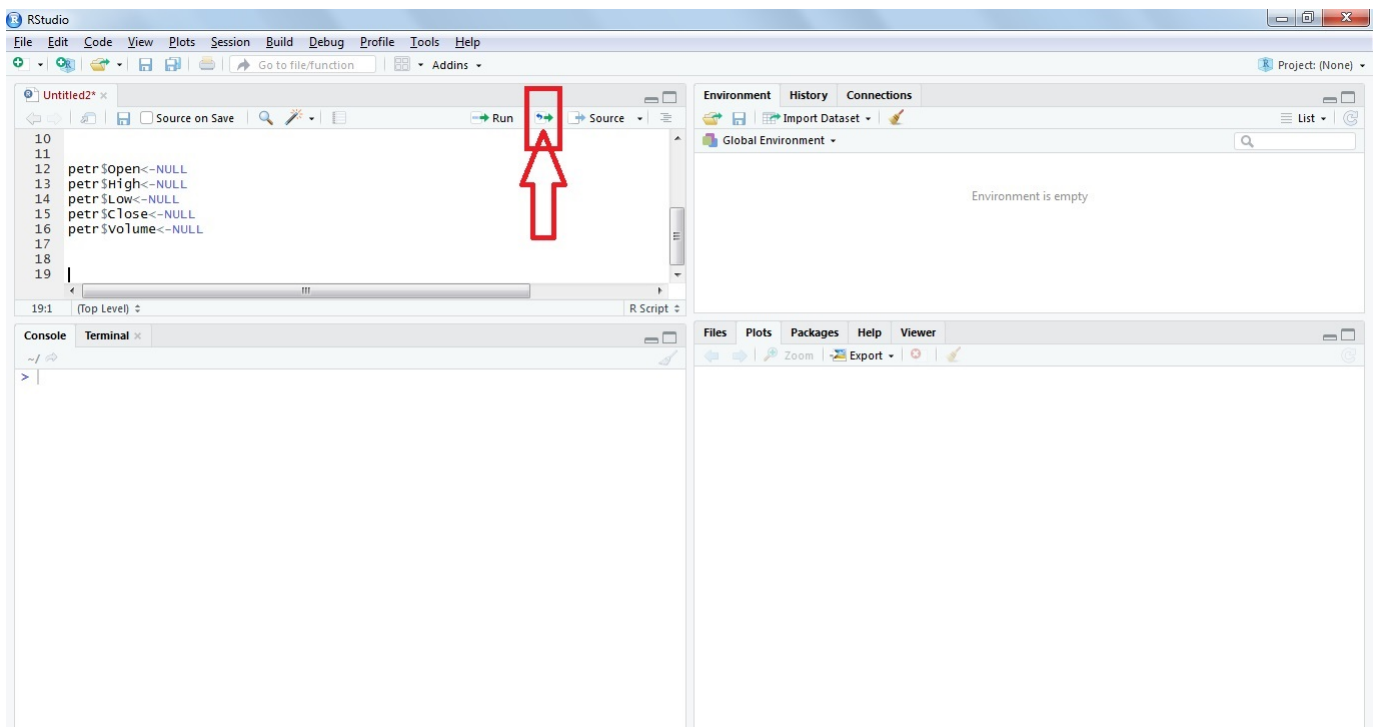


Figura 5.3.7: R Studio - Botão Re-run.

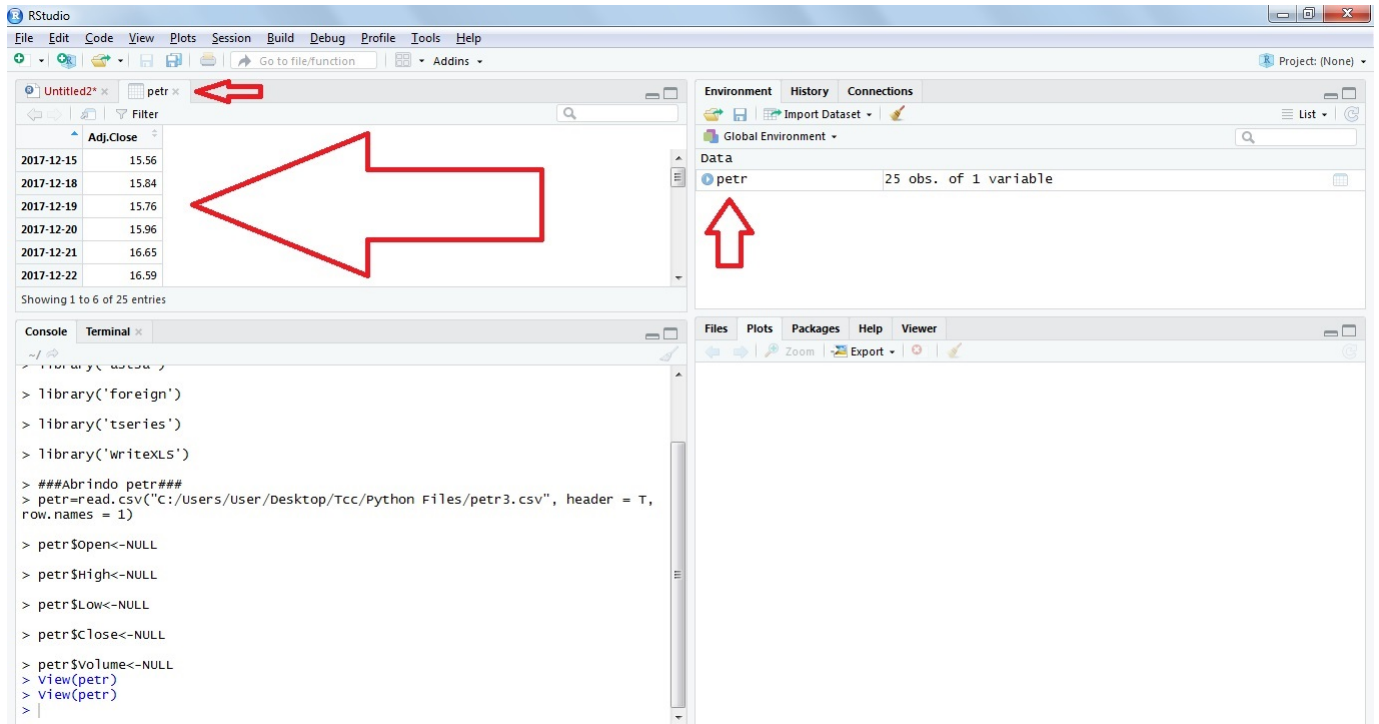


Figura 5.3.8: R Studio - Tabela de preços.

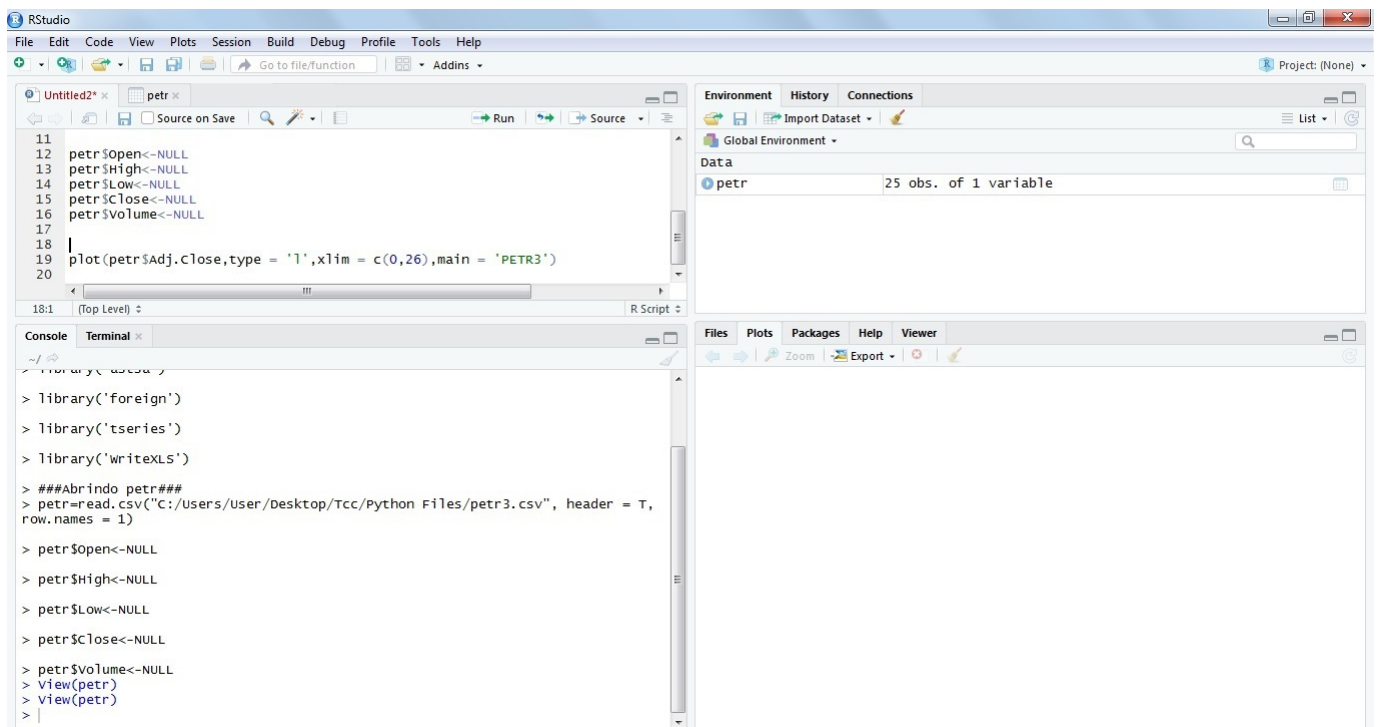


Figura 5.3.9: R Studio - Escrita do comando plot.

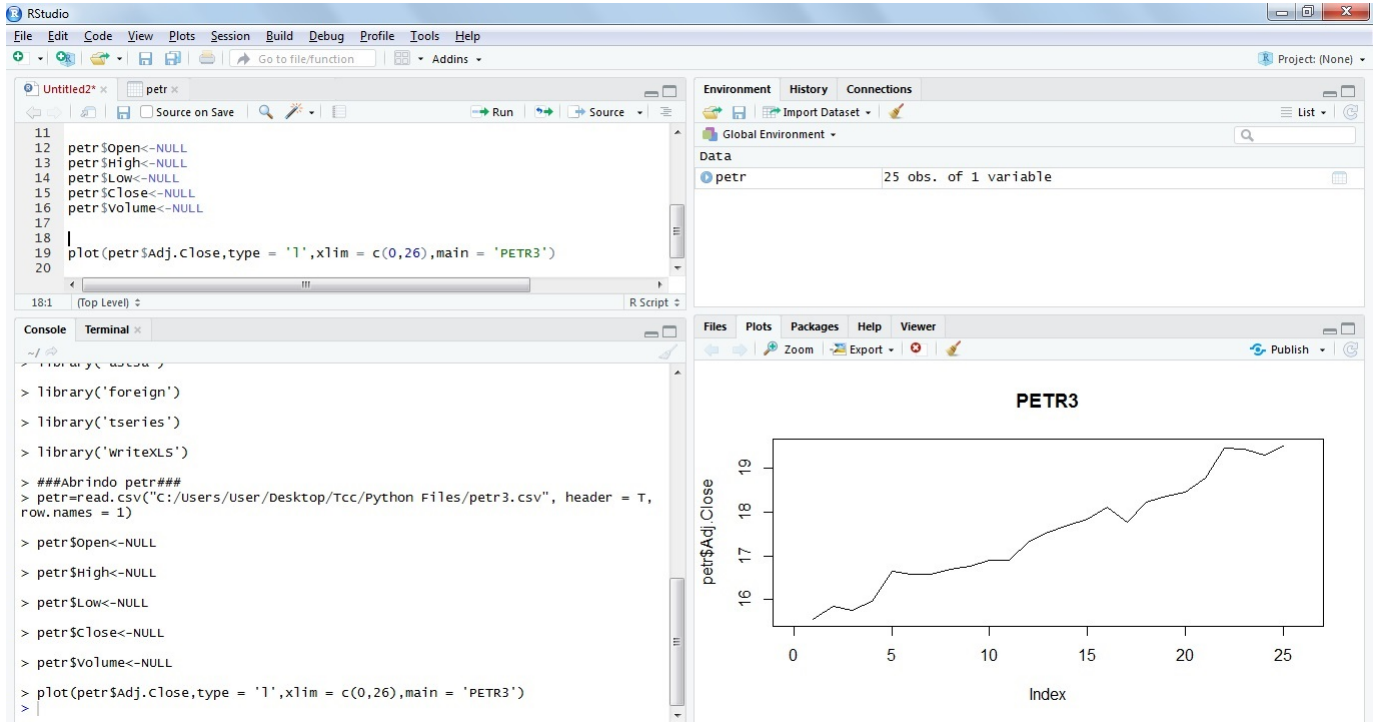


Figura 5.3.10: R Studio - O gráfico.

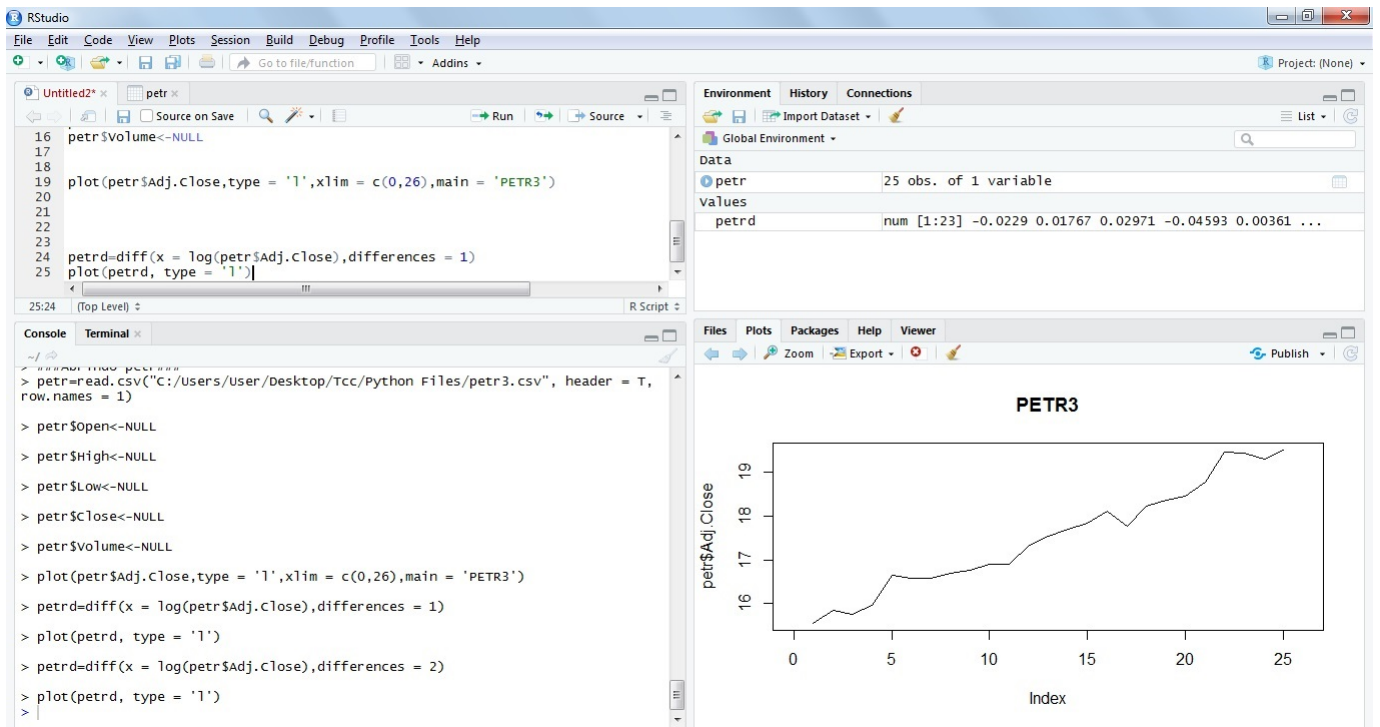


Figura 5.3.11: R Studio - Escrita do comando diff.

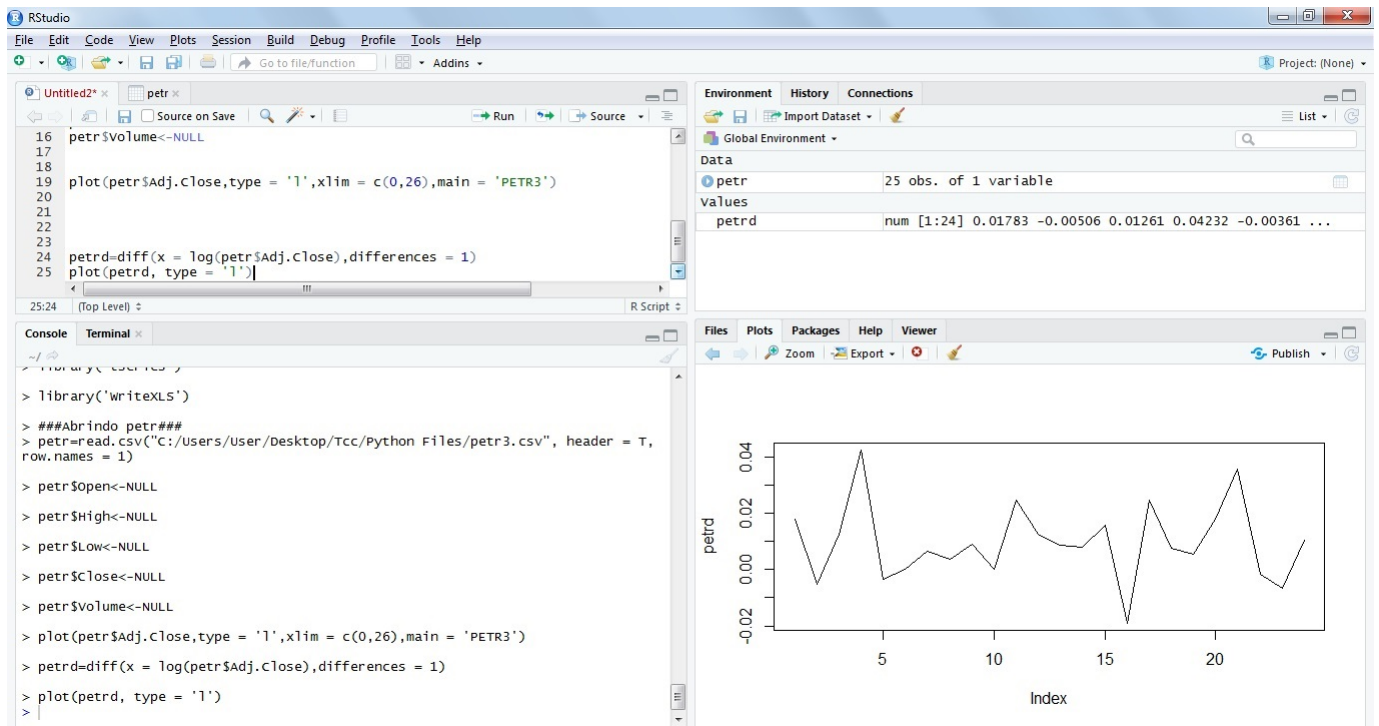


Figura 5.3.12: R Studio - O gráfico dos preços diferenciados 1 vez.

5.3.6 Plotando a FAC.

Agora que a série está estacionária podemos calcular a FAC. Para isso vamos utilizar o seguinte comando:

$$\text{acf}(x = \text{petr.d}, \text{main} = \text{'PETRDiferenciada'})$$

o subcomando “main” dá título ao gráfico. Após digitar o comando devemos clicar em “Re-run”. Acompanhe as figuras 5.3.13 e 5.3.14.

Assim, concluímos que a série tem memória em média igual à 1.

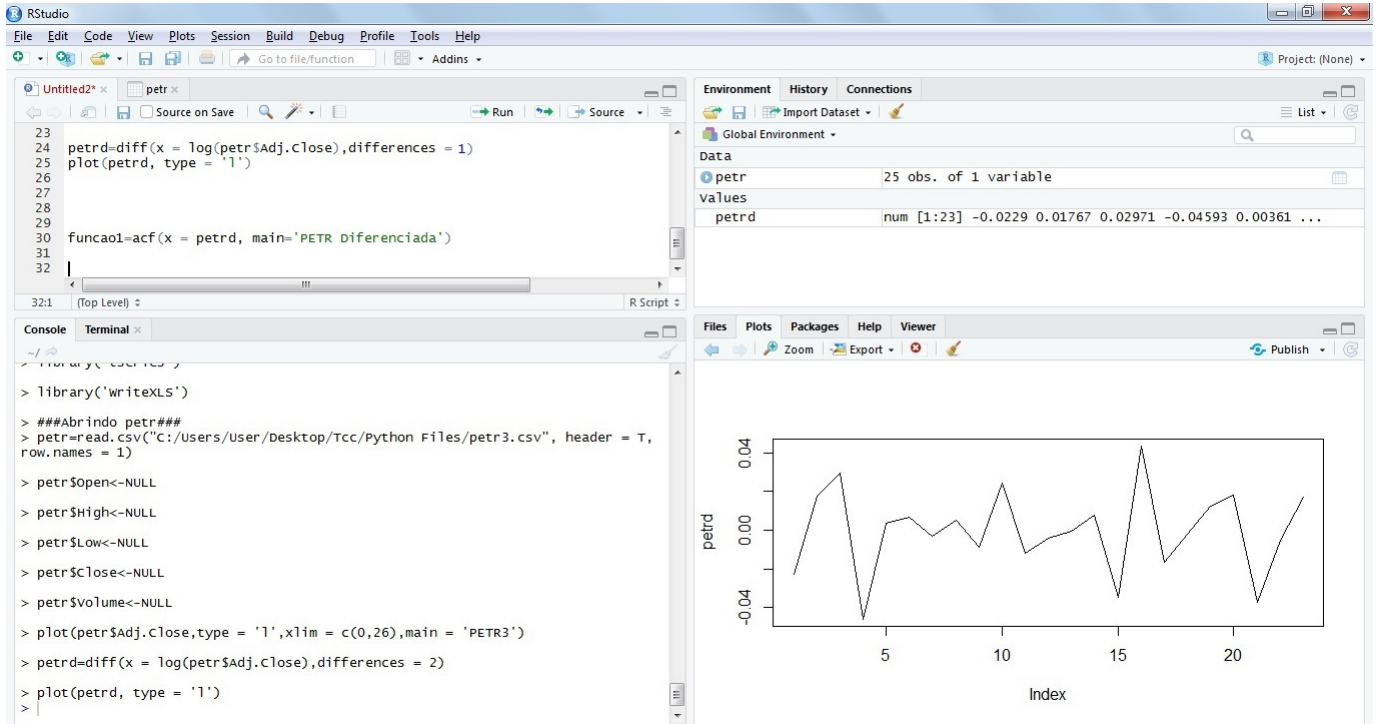


Figura 5.3.13: R Studio - Escrita do comando ACF.

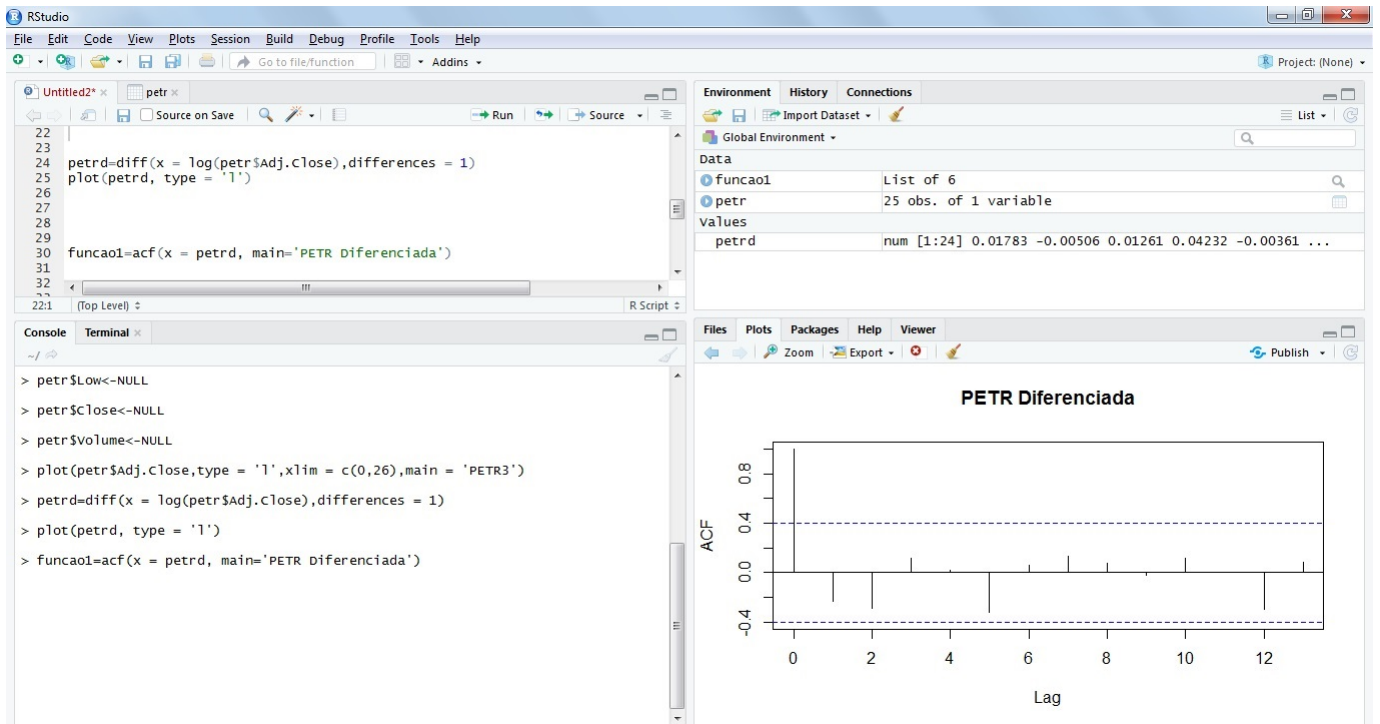


Figura 5.3.14: R Studio - O gráfico da FAC.

5.3.7 Plotando a FACP.

Como a série é estacionária podemos calcular a FACP do processo. Para isso vamos utilizar o seguinte comando:

```
pacf(x = petr, main = 'PETRDiferenciada')
```

Após digitar o comando devemos clicar em “Re-run”. Acompanhe as figuras 5.3.15 e 5.3.16.

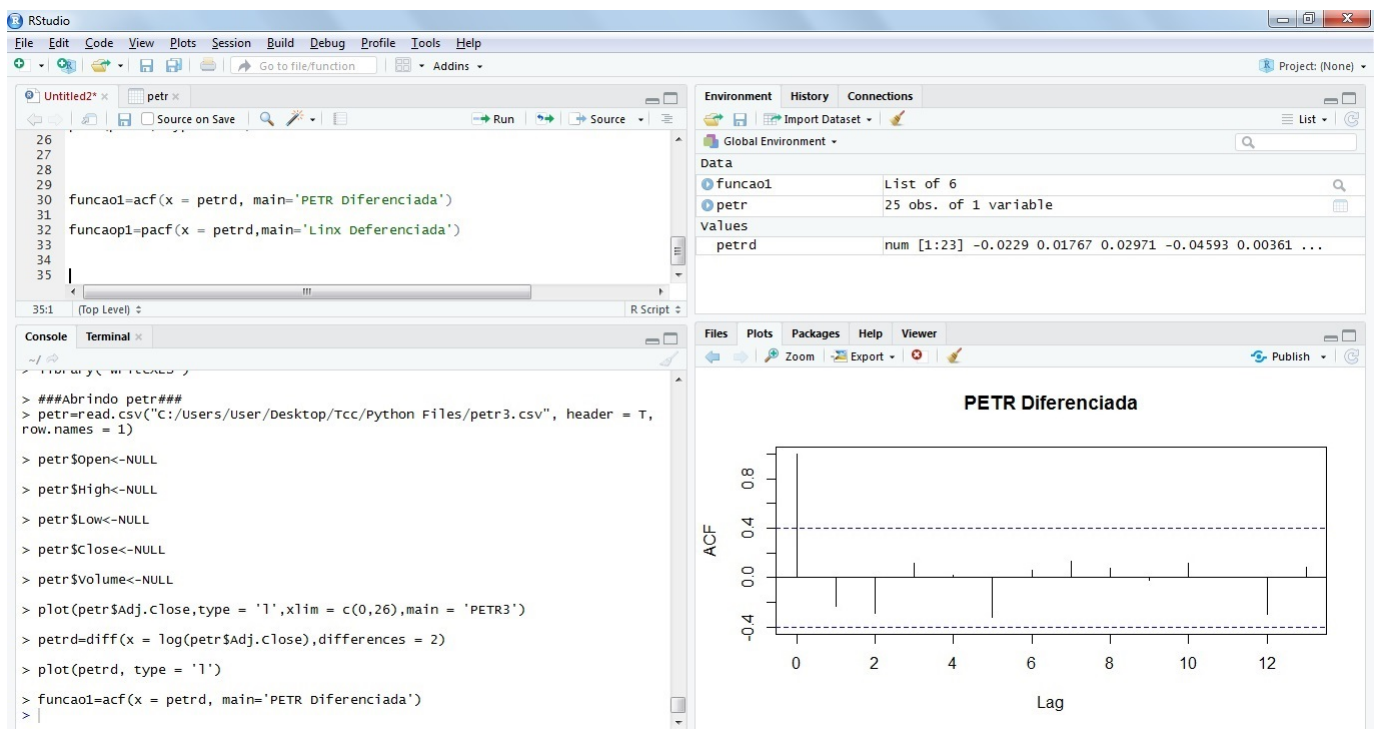


Figura 5.3.15: R Studio - Escrita do comando PACF.

Assim, concluímos que a série tem um índice de regressão igual à 1.

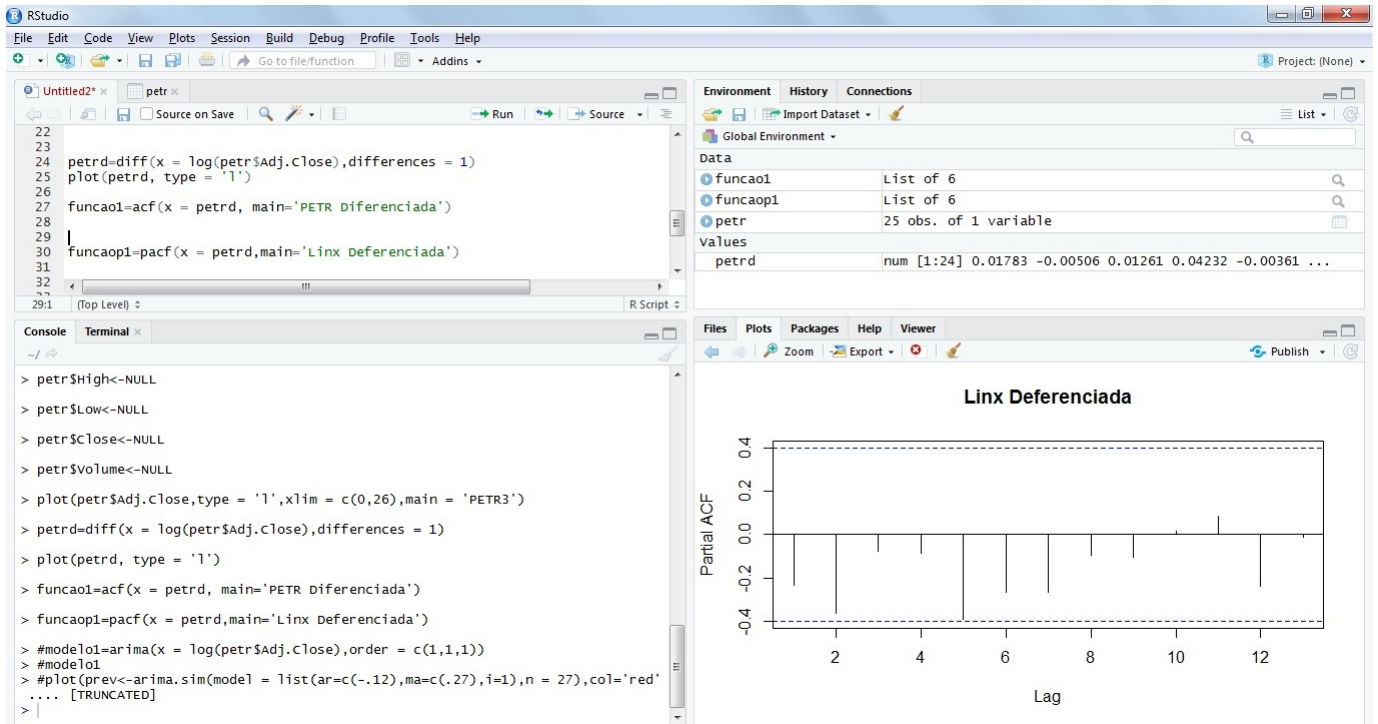


Figura 5.3.16: R Studio - O gráfico da FACP.

5.3.8 Os coeficientes do Modelo ARIMA.

Para especificar os coeficientes ϕ e θ modelo ARIMA(P,D,Q) utilizaremos o comando:

$$\text{arima}(x = \log(\text{petr}\$Adj.Close), \text{order} = c(1, 1, 1))$$

O subcomando “order = c(p,d,q)” é o comando onde devemos colocar os parâmetros para que o R studio faça a modelagem. Como observamos, o parâmetro P foi igual a 1 pela FACP, o parâmetro Q foi igual a 1 pela FAC e o parâmetro D foi igual a 1 pelo número de diferenciações. Após clicar em “Re-run” o programa nos forcencará os coeficientes ϕ e θ . Veja na figura 5.3.17.

Observamos que são coeficientes aceitos pelos estudos feitos, já que $|\phi|$ e $|\theta|$ são ambos menores que 1.

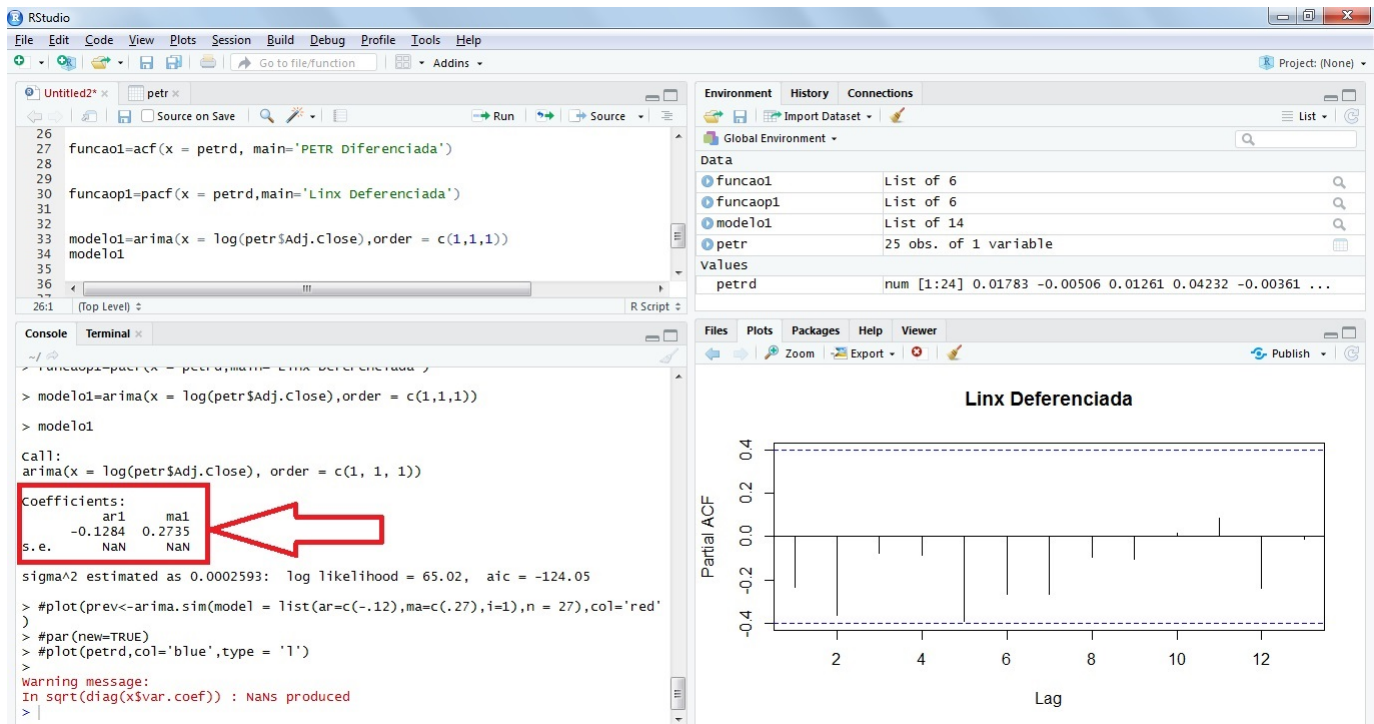


Figura 5.3.17: R Studio - Coeficientes ϕ e θ .

5.3.9 Plotando o modelo e comparando com o Original

Agora que temos os índices e seus respectivos coeficientes podemos rodar algumas simulações. Para isso vamos utilizar o comando:

```
arima.sim(model = list(ar = c(-.12), ma = c(.27), i = 1), n = 24), col = 'red'
```

O subcomando “n” define quantas amostras serão feitas, nesse caso foram 24 amostras. O subcomando “col” define a cor do gráfico, nesse caso os modelos serão plotados em cor vermelha e o original em cor azul. Assim basta plotar o modelo sobre o original utilizando entre os comandos “plot” o comando

```
par(new = TRUE)
```

. Acompanhe as figuras 5.3.18 e 5.3.19.

Podemos lançar quantos modelos quisermos, basta clicar em “Re-run” e um novo modelo será criado.

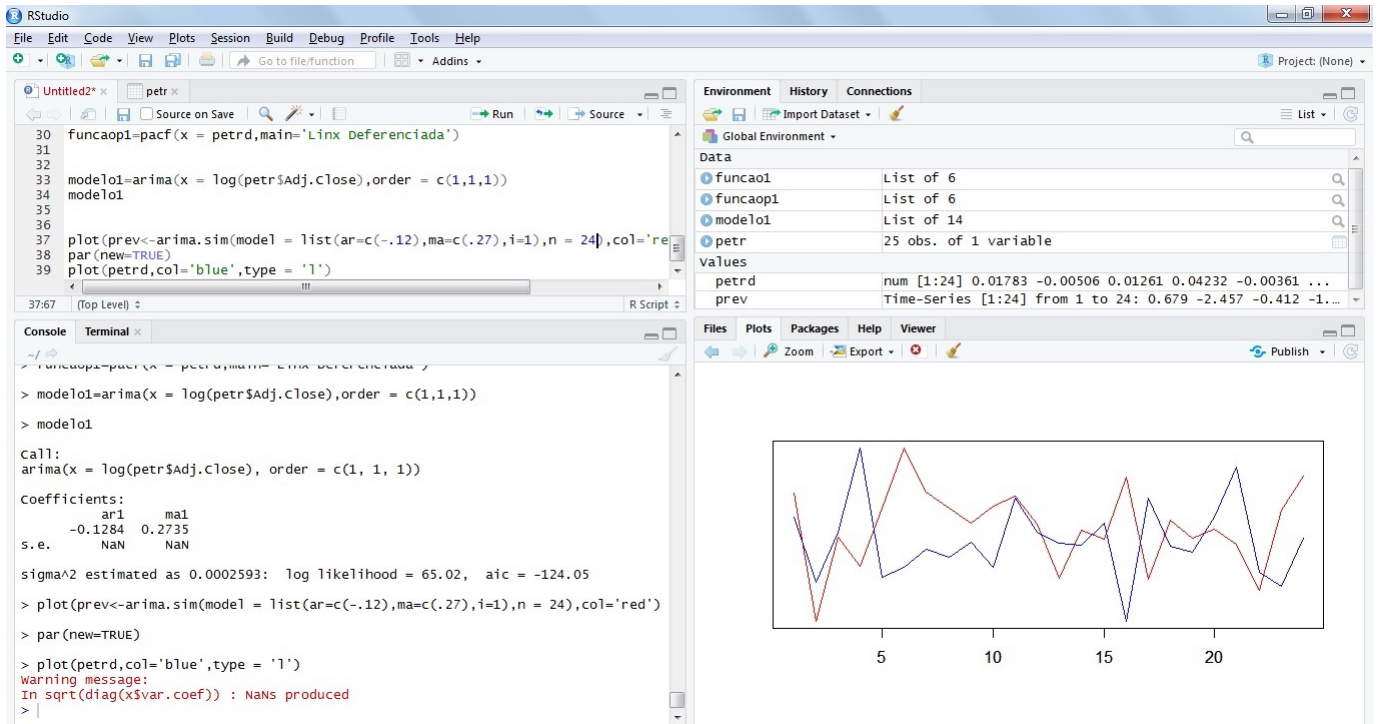


Figura 5.3.18: R Studio - modelo1.

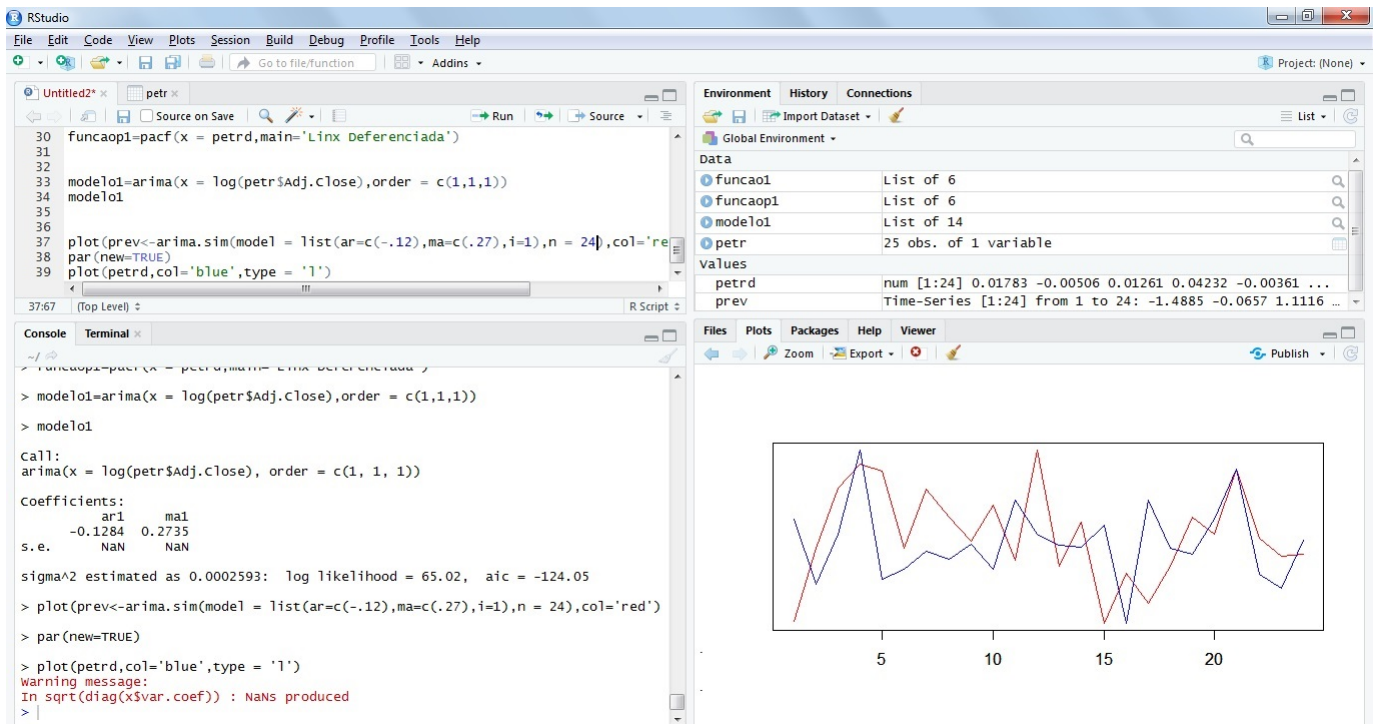


Figura 5.3.19: R Studio - modelo2.

CONCLUSÃO

Em nossa primeira etapa conseguimos motivar todos os alunos a compreenderem um pouco mais sobre porcentagens e suas aplicabilidades, pois de acordo com [8] não basta ao cidadão entender as porcentagens expostas em índices estatísticos, como o crescimento populacional, taxas de inflação ou desemprego, é preciso analisar/relacionar criticamente os dados apresentados, questionando/ponderando até mesmo sua veracidade. Assim como não é suficiente ao aluno desenvolver a capacidade de organizar e representar uma coleção de dados, faz-se necessário interpretar e comparar esses dados para tirar conclusões.

Sobre os cálculos, apenas 40% dos alunos conseguiram acompanhar. Contudo, a manipulação dos dados pelo “site”, no final do projeto, foi bem aceita por todos. Portanto, concluímos que a primeira etapa e o impacto social foram atingidos com sucesso, apesar de uma deficiência técnica dos alunos.

Ainda na linha de pensamento da obra [8], a competência nesses assuntos permite aos alunos uma sólida base para desenvolverem estudos futuros e atuarem em áreas científicas como a biologia e as ciências sociais. Assim, ao considerarmos o mundo em rápida mudança como o que estamos vivendo, é imprescindível o conhecimento da probabilidade de ocorrência de acontecimentos para agilizarmos a tomada de decisão e fazermos previsões.

Em vista disso, percebe-se a importância para a Ciência, no geral, de mantê-los motivados. Uma continuidade do trabalho na parte educacional seria um plano pedagógico para que os alunos aprendessem programação em R Studio com planilhas.

Conseguimos fazer simulações de um modelo ARIMA (2,2,1) que se assemelhava ao comportamento da série de preços de fechamento diário das ações da Linx, LINX3, quando diferenciadas 2 vezes. Durante os estudos percebemos que os modelos AR,

MA, ARMA e ARIMA são suficientes para acompanhar alguns movimentos de ativos do mercado financeiro, porém uma dificuldade é encontrada quando a escala de tempo é diminuída de dias para minutos. O tempo que demoramos para copiar os dados e analisar qual seria o melhor modelo é maior que o tempo real dos acontecimentos. Uma continuidade do trabalho na parte técnica seria a criação de um programa em Python que informa, em tempo real, se o gráfico intradiário de um ativo da Bolsa de valores, B3, apresenta um comportamento do tipo: AR, MA, ARMA ou ARIMA.

O estudo desses modelos chamou a atenção de professores e também da direção da escola para o passo anterior: os juros compostos. Isso é positivamente impactante, pois muitos professores não tinham clareza no poder dos juros compostos e com a demonstração do interesse de outros professores, os alunos ficaram ainda mais motivados. Houve o apoio da direção escolar ao reservar, por um semestre, a Sala Multimídia com internet. Mesmo com uma deficiência técnica os alunos que já trabalham perceberam o quanto poderiam mudar o seu futuro com apenas R\$ 200,00 por mês. Muitos alunos, durante os estudos, já começaram a guardar o seu dinheiro, investindo em Tesouro Direto e aprendendo a lidar com a inflação.

Por fim, após a aplicação das aulas teóricas e atividades, acreditamos que o impacto nos alunos de baixa renda tenha sido atingido pela fomentação do ensino da matemática financeira através da modelagem Matemática e da Modelagem Computacional construindo um caminho viável para uma maior consciência financeira do aluno.

BIBLIOGRAFIA

- [1] BASES LEGAIS-Parâmetros Curriculares Nacionais BRASIL and Parâmetros Curriculares Nacionais Ensino Médio. Pcn+ para o ensino de ciências e matemática. *Brasília: Ministério da Educação, 2002.*
- [2] Eduardo de Brito. Problematização no ensino médio: Como lucrar na bolsa de valores. *Universidade de Federal do ABC, 2014.*
- [3] Artur de Oliveira Lopes & Sílvia Lopes. *Introdução aos Processos Estocásticos para estudantes de Matemática.* UFRGS, 2016.
- [4] Fritz Moritz Heichelheim. *An ancient economic history: from the palaeolithic age to the migrations of the Germanic, Slavic and Arabic nations*, volume 1. AW Sijthoff, 1958.
- [5] EVES Howard. Introdução à história da matemática. *Tradução Hygino H. Domingues, 3, 1995.*
- [6] Barry R James. *Probabilidade: um curso em nível intermediário*, volume 12. Instituto de Matemática Pura e Aplicada, CNPq, 1981.
- [7] Andreï Nikolaevich Kolmogorov. Foundations of the theory of probability. *University of Oregon, 1956.*
- [8] Celi Espasandin Lopes. O ensino da estatística e da probabilidade na educação básica e a formação dos professores. *Cad. Cedes, Campinas, 28(74):57–73, 2008.*
- [9] Pedro A Morettin. *Econometria financeira. Um curso em séries temporais, 2008.*
- [10] Gelson Iezzi & Carlos Murakami. *Fundamentos de matemática elementar, 1: conjuntos, funções.* Atual, 1993.
- [11] A de L PUCCINI. *Matemática financeira e análise de investimentos.* Ed. Beta, 1977.
- [12] Pabli Martín Rodríguez. *Passeios aleatórios.* Universidade de São Paulo, 2014.
- [13] Sidney Homer & Richard Sylla. *A History of Interest Rates*, volume 322. John Wiley & Sons, 2005.

- [14] Pedro A. Morettin & Clélia Tolo. *Análise de séries temporais*. Blucher, 2006.
- [15] Otaviano Augusto Marcondes Helene & Vito R. Vanin. *Tratamento estatístico de dados em física experimental*. E. Blucher, 1981.



APÊNDICE A - QUESTIONÁRIO

QUESTIONÁRIO SOBRE O PROJETO DE FINANÇAS:

Nome: _____

Idade: _____

Escolaridade: _____

Sexo: F () M ()

1. Você guarda dinheiro?

() Sim.

() Não.

2. Quanto?

() Não guardo.

() Menos de R\$100.

() Entre R\$100 e R\$200.

() Entre R\$200 e R\$400.

() Mais de R\$400.

3. O que o dinheiro representa para você?

4. Você acredita que é possível ser um milionário antes de obter a aposentaria oferecida pelo governo? Se sim, qual a sua estratégia?

5. Cite os tipos de investimentos que você conhece.

6. Você sabe que existe diferença entre Juros Simples e Juros Compostos?

Sim.

Não.

7. Você consegue identificar a propriedade (função) matemática dos Juros Simples e dos Juros Compostos? Se sim, escreva.

Observou-se que as pessoas com escolaridade maior têm uma tendência de guardar mais dinheiro. Observou-se que as pessoas com escolaridade menor tem tendência a responder *não* à pergunta 6.

B

APÊNDICE B - PROGRAMAÇÃO PYTHON DE MOVIMENTO BROWNIANO

```
from numpy.random import standard_normal
from numpy import array, zeros, sqrt, shape
from pylab import *
```

```
preco_inicial = 10 (preço inicial da ação)
distancia = 1 (parametro para o calculo da quantidade de passos)
tamanho_do_passo = 0.005 (tamanho do passo)
```

(Por exemplo; distancia=1 e tamanho_do_passo = 0.005 teremos 200 passos)

```
sigma = 0.03 (volatilidade diária, no caso 3%)
```

```
mu = 0.01 (retono médio de cada passo)
```

```
nSimu = 3 (número de simulações)
```

```
passos=round(distancia/tamanho_do_passo) (passos em dias)
S = zeros([nSimu, passos], dtype=float)
x = range(0, int(passos), 1)
```

```
for j in range(0, nSimu, 1):
```

```
    S[j,0]= preco_inicial
```

```
    for i in x[:-1]:
```

```
        S[j,i+1]=S[j,i]+S[j,i]*(mu-0.5*sigma*sigma)*tamanho_do_passo+sigma*S[j,i]
```

```
*sqrt(tamanho_do_passo)*standard_normal();  
plot(x, S[j])  
  
title('Simulacao com %d dias e Preco Inicial de %.2f' % (int(passos), tamanho_do_passo))  
xlabel('Dias')  
ylabel('Preço da Acao')  
show()
```

O resultado são gráficos como os da figura 4.4.1.

C

APÊNDICE C - PROGRAMAÇÃO EM R STUDIO DE MODELO ARIMA

```
##### bibliotecas #####

library('forecast')
library('astsa')
library('foreign')
library('tseries')
library('WriteXLS')

##### abrindo o arquivo csv e eliminando as colunas desnecessárias #####

linx=read.csv("C:/linx.csv", header = T, row.names = 1)
linx$Open<-NULL
linx$High<-NULL
linx$Low<-NULL
linx$Close<-NULL
linx$Volume<-NULL

##### gráfico e diferenças #####

plot(linx$Adj.Close,type = 'l',xlim = c(0,30),main = 'LINX3',sub = '4/11/2015 - 15-12-2015',xlab = 'Número de Observações',ylab = 'Fechamento')
linxd=diff(x = log(linx$Adj.Close),differences = 1)
plot(linxd,type = 'l',main = 'LINX3 DIFERENCIADA 1 VEZ',xlab = 'observações',ylab = 'Diferença')
linxd=diff(x = log(linx$Adj.Close),differences = 2)
plot(linxd,type = 'l',main = 'LINX3 DIFERENCIADA 2 VEZES',xlab = 'observações',ylab
```

```
= 'Diferença')
```

```
##### Autocorrelação #####
```

```
funcao=acf(x = linx$Adj.Close,main='Linx Fechamento Original')  
funcao1=acf(x = linxd, main='Linx Diferenciada')
```

```
##### Autocorrelação Parcial #####
```

```
funcaop=pacf(x = linx$Adj.Close,main='Linx Fechamento Original')  
funcao1=pacf(x = linxd,main='Linx Diferenciada')
```

```
##### arima #####
```

```
modelo1=arima(x = linxd,order = c(2,0,1))  
modelo2=arima(x = linx$Adj.Close,order = c(2,2,1))
```

```
##### gráficos de verificação #####
```

```
plot(arima.sim(model = list(ar=c(-.92,-.76),ma=c(.18)),n = 27),col='red')  
par(new=TRUE)  
plot(linxd,col='blue',type = 'l')
```