

UNIVERSIDADE FEDERAL DE ALAGOAS
Instituto de Física

Frederico Salgueiro Passos

O Teorema das Seções de Lévy Aplicado à
Séries Temporais Correlacionadas Não
Estacionárias: Uma Análise da Convergência
Gaussiana em Sistemas Dinâmicos

Maceió - AL
2014

FREDERICO SALGUEIRO PASSOS

O Teorema das Seções de Lévy Aplicado à
Séries Temporais Correlacionadas Não
Estacionárias: Uma Análise da Convergência
Gaussiana em Sistemas Dinâmicos

Tese apresentada ao
Instituto de Física da Universidade
Federal de Alagoas como
pré-requisito para obtenção
do título de Doutor em Ciências

Orientador: Prof. Dr. Iram Marcelo Gléria
Co-orientador: Prof. Dr. César Moura Nascimento

Maceió - AL
2014

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca Central

Bibliotecária Responsável: Helena Cristina Pimentel do Vale

P285t Passos, Frederico Salgueiro.
O teorema das seções de Lévy aplicado à séries temporais correlacionadas não estacionárias: uma análise da convergência gaussiana em sistemas dinâmicos / Frederico Salgueiro Passos. - 2014.
125 f. : il.

Orientador: Iram Marcelo Gléria.
Coorientador: César Moura Nascimento.
Tese (Doutorado em Física da Matéria Condensada) – Universidade Federal de Alagoas. Instituto de Física. Maceió, 2014.

Bibliografia: f. 121-125.

1. Teorema das seções de Lévy. 2. Teorema do limite central. 3. Séries Temporais. 4. Heteroscedásticas. 5. Testes de convergência. 6. Física estatística.
I. Título.

CDU: CDU: 53:519.246.8

PARECER DA BANCA EXAMINADORA DE DEFESA DE
TESE DE DOUTORADO

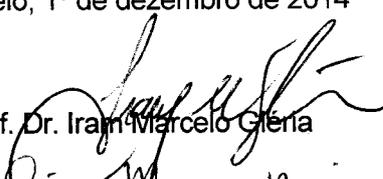
“O Teorema das Seções de Lévy aplicado a Séries temporais correlacionadas não estacionárias: Uma análise da convergência gaussiana em sistemas dinâmicos”

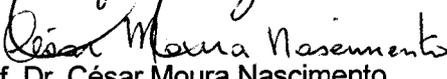
por

Frederico Salgueiro Passos

A Banca Examinadora composta pelos professores Iram Marcelo Gléria (orientador), do Instituto de Física da Universidade Federal de Alagoas, César Moura Nascimento, (co-orientador), do Departamento de Física da Universidade Federal de Sergipe, Askery Alexandre Canabarro Barbosa da Silva, do curso de Física da Universidade Federal de Alagoas (Campus Arapiraca), Samuel Silva de Albuquerque, do curso de Física da Universidade Federal de Alagoas (Campus Arapiraca), Rodrigo de Paula Almeida Lima, do Instituto de Física da Universidade Federal de Alagoas, Vinicius Manzoni Vieira, do Instituto de Física da Universidade Federal de Alagoas, Carlos Argolo Pereira Alves, do Instituto Federal de Alagoas consideraram o candidato **aprovado** com grau “ A ”.

Maceió, 1º de dezembro de 2014


Prof. Dr. Iram Marcelo Gléria


Prof. Dr. César Moura Nascimento


Prof. Dr. Askery Alexandre Canabarro Barbosa da Silva


Prof. Dr. Samuel Silva de Albuquerque


Prof. Dr. Rodrigo de Paula Almeida Lima


Prof. Dr. Vinicius Manzoni Vieira


Prof. Dr. Carlos Argolo Pereira Alves

DEDICATÓRIA

Dedico à minha esposa, meus pais e meu irmão que tornaram esse sonho possível

AGRADECIMENTOS

Muitos nomes devem ser citados e muitos outros podem ser esquecidos nessa lista de agradecimentos, agradeço então inicialmente a todos que de alguma forma participaram na construção dessa tese, mesmo que seus nomes não estejam aqui, saibam fazem parte do todo.

O começo de tudo se deve a minha família, agradeço aos meus pais, Newton e Denise e ao meu irmão Rodolfo, por toda ajuda e apoio.

Com certeza devo agradecer aos meus grandes amigos que me acompanharam durante toda essa jornada do conhecimento, Alex Costa, Anderson, Henrique, Neto e especialmente Tiago Lobo, que travou comigo batalhas nunca antes vistas, apenas para ver que eu estava certo.

Embora tenhamos pouco contato agora, gostaria de agradecer a todos os professores e funcionários do IF-UFAL por me ensinarem as bases que hoje eu monto a minha tese, sem o apoio e troca de experiências, essa tese não teria a robustez que hoje apresenta.

Meus eternos agradecimentos ao Professor Madras Viswanathan Gandhimohan, que foi o meu primeiro orientador e hoje tenho o puro orgulho de poder dizer que além de um colaborador para os meus projetos pessoais é meu amigo. Sem suas palavras de incentivo e de “fome” pelo conhecimento, não teria vislumbrado tanto.

Xadrez é um jogo que exige paciência e concentração, assim como o kendô. Agradeço aos meus amigos de treino que me ajudaram a criar estratégias e manter firmes as bases quando tempos difíceis chegam. Agradeço a Renato Simões, Luiz Gustavo Vasques, Tadeu Patêlo, Ranilson Paiva e Felipe Dória, que não treina mais conosco mas foi um dos meus principais incentivadores.

Quero também agradecer aos meus colegas de trabalho no IFAL campus Marechal Deodoro, que me deram todo o apoio para concluir do doutorado e me incentivaram ao máximo para criar novas linhas de pesquisa e me abraçaram como uma família. Agradeço a José Ginaldo Júnior, Renato Romero, Pedro Guilherme e Eder de Souza por toda a força.

Também quero agradecer em especial aos meus avôs que foram pilares sólidos e todas as hipóteses sempre apoiaram qualquer sonho louco que eu poderia ter, desde tecladista até mesmo querer ser o homem-aranha. Seus ensinamentos e palavras transcendem as barreiras da vida e se mantêm como velas acesas na escuridão me guiando. Agradeço imensuravelmente a eles todos, Álvaro, Zilah, Newton e Célia, obrigado pelos sonhos.

De todo o coração gostaria de agradecer a toda a família Amorim e seus agregados que me receberam como parte de sua família, Bruno, Vanessa, Rafael, Letícia, Biu, Lourdes, Amauri e Valmira.

Eu gostaria também de agradecer a todos os meus grandes amigos que estão comigo desde os primórdios da minha criação, Eli Júnior, Guilherme, Pedro Ivo, José Neto, Tia Lúcia, Tio Eli, Tio Fernando, Tia Vivian, Tia Alessandra, Tio Toninho e Diogo Braga.

Venho também agradecer a todos os membros da minha família, que mesmo distantes sempre se demonstraram interessados nas maluquices que eu falava e tentava explicar. São muitos nomes para serem colocados mas quero agradecer uns em especial, aos meus tios maternos, Alexandre Sérgio, Marco Aurélio e Guilherme Salgueiro, que são exemplos como tios e como família, pessoas com as quais eu posso contar a qualquer dia e qualquer hora independente do que haja. Minha tia paterna Nilcéia e sua família, e a minha tia/prima Márcia que é um exemplo de profissional na qual me espelho.

Especialmente, gostaria de agradecer a minha esposa Evelyn Amorim Dias Passos, minha companheira, amiga, namorada, amante, ajudante, sonhadora, inquieta, deslumbrante e acima de tudo tolinha. Uma das principais pessoas a me apoiar(e ordenar) na conclusão deste trabalho e um porto seguro para nossos planos tão amplos para o futuro. Agradecer é pouco para alguém que está tão ao meu lado, e estará sempre e até depois. MocemxqtdeveS2.

Sem esquecer, gostaria de agradecer as agencias de financiamento que proporcionaram o apoio necessário para produzir a tese. Agradeço ao CNPq, a CAPES, a FAPITEC/CNPq e ao IFAL que financia parte desse trabalho através de uma bolsa de produtividade em pesquisa.

2. ((:

“EPPUR SI MOUVE!”
(tradução livre: "no entanto ela se move")
-Galileu Galilei, 1633

RESUMO

Processos não-estacionários com interações fracas aparecem como problemas desafiadores em sistemas complexos em física. Uma questão interessante é como quantificar a taxa de convergência para o comportamento gaussiano em séries temporais heteroscedásticas, sem uma variância única em toda a série, provenientes de sistemas financeiros, reescaladas com os primeiros momentos estacionários mas com uma multifractalidade não estacionária e segundos momentos que possuem uma correlação do longo alcance e verificar o mesmo mecanismo também em séries temporais geradas a partir de um movimento Browniano Fracionado onde a correlação da série depende de um parâmetro ajustável. Aqui é usada uma extensão do teorema das seções de Lévy. Analisando as propriedades estatísticas e multifractais de uma série temporal heteroscedástica e encontrando que as seções de Lévy fornece uma convergência mais rápida para o comportamento gaussiano relativo à convergência das tradicionais somas de variáveis, o teorema do limite central. Para entender essa transição foram utilizados vários testes estatísticos que forneceram dados suficientes sobre o comportamento de convergência. Também observou-se que os sinais reescalados mantêm suas propriedades multifractais mesmo depois de atingirem um regime que parece ser um regime gaussiano.

Palavras-Chave: Teorema das Seções de Lévy. Teorema do Limite Central. Séries Temporais Heteroscedásticas. Testes de Convergência. Física Estatística .

ABSTRACT

Weakly nonstationary processes appear in many challenging problems related to the physics of complex systems. An interesting question is how to quantify the rate of convergence to Gaussian behavior of rescaled heteroscedastic comming from economics time series with stationary first moments but nonstationary multifractal long-range correlated second moments and also time series generated from fractionated brownian motion where the series correlation is dependent of a parameter. Here it is used the approach which uses a recently proposed extension of the Lévy sections theorem. It was analyzed the statistical and multifractal properties of heteroscedastic time series and found that the Lévy sections approach provides a faster convergence to Gaussian behavior relative to the convergence of traditional partial sums of variables. To understand this transition it is used several statistical tests to provide enough data on convergence behavior. It was also observed that the rescaled signals retain multifractal properties even after reaching what appears to be the stable Gaussian regime.

Keywords:Lévy Sections Theorem. Central Limit Theorem. Heteroscedastic Time Series. Convergence Tests. Statistical Physics.

LISTA DE FIGURAS

2.1	Função Densidade de Probabilidade de um dado viciado segundo a tabela 2.1	27
2.2	Função Densidade de Probabilidade do oscilador harmônico quântico, com parâmetros livres.	28
2.3	Função Densidade de Probabilidade da função descrita na equação 2.11 onde são indicados os valores da média, mediana e moda, podendo identificar as características de cada uma delas, especialmente a moda onde claramente é o ponto mais alto da probabilidade. A mediana divide a distribuição em duas partes iguais, ou seja, a soma das probabilidades de os valores possíveis da variável de $-\infty$ até 1, é igual a soma de todas as probabilidades dos valores acima de 1 até ∞ . E a média, calculada pela equação 2.10, é a tendência, ou centro de massa, da distribuição.	30
2.4	Nessa figura são apresentadas algumas distribuições que possuem variados valores de <i>skewness</i> , podendo mostrar suas propriedades assimétricas. (a) Distribuição com <i>skewness</i> positiva, segundo a equação 2.11, mostrando uma assimetria à direita. (b) Distribuição com <i>skewness</i> negativa, do espelho da função dada pela equação 2.11 com $x = -x$, mostrando uma assimetria à esquerda. (c) Distribuição com <i>skewness</i> nula, da distribuição normal, completamente simétrica sobre o eixo da média.	34
2.5	Diferentes distribuições com variados valores de <i>kurtosis</i> . (a) Distribuição com <i>kurtosis</i> excessiva positiva, segundo a distribuição de Laplace, com <i>kurtosis</i> excessiva igual a 3. (b) Distribuição com <i>kurtosis</i> excessiva negativa, segundo a distribuição de Wigner, com <i>kurtosis</i> excessiva igual a -1. (c) Distribuição com <i>kurtosis</i> excessiva nula, a distribuição gaussiana.	35
2.6	Posição x de caminhantes aleatórios com passos de tamanho $l = 1$ e com probabilidades: (a) 0.5, (b) 0.45 e (c) 0.55, em função do número de passos dados.	39
2.7	Probabilidade $P_X(x)$ de caminhantes aleatórios com passos de tamanho $l = 1$ e com probabilidades p : (a) 0.5, (b) 0.45 e (c) 0.55, versus a posição x	40
2.8	Alguns exemplos de curvas da distribuição gaussiana com alguns parâmetros diversos. É possível notar que a distribuição é simétrica, possuindo portanto <i>skewness</i> nula e tem um formato bastante característico, em forma de sino.	42
2.9	Algunas distribuições provenientes da equação 2.95 para diferentes α e com $\beta = \mu = 0$ e com $c = 1$. Dados gerados usando o programa R[20]	48
2.10	Algunas distribuições provenientes da equação 2.95 com $\mu = 0$, $c = 1$, (a) $\alpha = 0.5$, (b) $\alpha = 1.0$, (c) $\alpha = 1.5$, (d) $\alpha = 1.75$ e variados valores de β . Dados gerados usando o programa R[20]	48

LISTA DE FIGURAS

2.11	Eixo em semi-log da figura 2.9, mostrando como o decaimento sugere uma lei de potência, indicado pelo comportamento não polinomial das distribuições quando sujeitas ao logaritmo da probabilidade, <i>i.e.</i> $\ln e^{-x^2} = -x^2$	49
2.12	Figura voltada à cauda, onde o comportamento em lei de potência para $\alpha < 2$ se torna mais evidente e observa-se a discrepância para $\alpha = 2$ quanto mais x se afasta do centro da distribuição.	49
2.13	A distribuição de Lévy-Smirnov para alguns parâmetros	52
2.14	A distribuição de Cauchy para alguns parâmetros	54
2.15	Representação ilustrativa da auto-similaridade presente em figuras geométricas e séries temporais. Adaptada de [25]	55
2.16	Exemplo do comportamento de $f(\alpha)$ em relação a α . Adaptada de [34]	58
3.1	Distribuição discreta de probabilidades para o lançamento de um dado justo. Todos os valores possíveis de se obter, os números naturais de 1 à 6, tem a mesma probabilidade de ocorrência, $1/6$	65
3.2	Curvas representando as diversas distribuições para a soma das faces voltadas para cima em conjuntos de lançamentos de dados	66
3.3	Distribuição gaussiana e sua FDA pra uma variável aleatória	74
3.4	Diversas distribuições gaussianas com diferentes parâmetros e suas FDA afim de comparação e observação do comportamento tomado por cada uma das funções acumuladas.	74
3.5	A Função de Distribuição Acumulada Empírica para um conjunto de dados	75
3.6	Comparação entre $F(x)$ e $F_n(x)$	76
3.7	Comparação entre $F_{1,n}(x)$ e $F_{2,n'}(x)$	77
4.1	Auto-correlação de retornos do S&P500 em um gráfico semi-log onde é possível observar o comportamento exponencial($e^{-\tau}$) com coeficiente de decaimento da ordem de $\tau = 4$ min. Adaptada de [54]	81
4.2	Auto-correlação de volatilidade do S&P500, decaindo com uma lei de potência do tipo $t^{-0,3}$. Adaptada de [54]	82
4.3	Distribuição de probabilidade para os retornos definidos como Z do S&P500, onde pode-se notar a presença de caudas grossas, ou seja, probabilidades não nulas de eventos extremos e a curva gaussiana(em pontos) não reflete o comportamento dos dados. Adaptada de [54]	82
4.4	(a)Dependência do expoente de Hurst com o q-ésimo momento. (b) Espectro multifractal em função do expoente de α , que no caso é conhecido como expoente de Hölder.[40]	83
5.1	Série de retornos do câmbio do DEM/USD fornecida por Olsen & Associates	85
5.2	Densidade dos retornos da série	86
5.3	Densidade dos retornos da série reduzida para 10^4 dados	86
5.4	Auto-correlação das séries de retorno	87
5.5	Auto-correlação das séries de volatilidade	87
5.6	Dependência do expoente de Hürst com o q-ésimo momento pelo método MF-DFA	88
5.7	Comportamento do espectro multifractal, calculado com as equações 2.138 e 2.139	88

5.8	Condição de Lindeberg para a série empírica do DEM/USD, mostrando que a série tem uma variância finita e por esse motivo os teoremas podem ser aplicados à ela.	89
5.9	As curvas de convergência da <i>kurtosis</i> para os dados analisados	90
5.10	As curvas de convergência da <i>skewness</i> para os dados analisados	91
5.11	Séries seguindo TLC antes e depois do regime de convergência	92
5.12	Zoom na figura 5.11 para destacar o D_n	92
5.13	Séries seguindo TSL antes e depois do regime de convergência	93
5.14	Zoom na figura 5.13 para destacar o D_n	93
5.15	Curvas de D_n versus τ para os dois teoremas	94
5.16	Curvas de D_{KL} versus τ para os dois teoremas	95
5.17	Teste KS comparando a série original do DEM/USD com a aplicação dos Teoremas.	96
5.18	Teste KL comparando a série original do DEM/USD com a aplicação dos Teoremas.	97
5.19	Expoente de Hürst médio em função do momento q . (a) Média para $\tau \leq 150$; (b) Média para $\tau > 150$	98
5.20	Espectro multifractal, $f(\alpha)$, em função de α . (a) Média para $\tau \leq 150$; (b) Média para $\tau > 150$	99
6.1	Diagrama de fase no plano (E/t,a), mostrando os estados estendidos e localizados.	102
6.2	Energias por sítios que foram geradas pela equação 6.9, com $N=2000$; onde $a = 0$ é uma sequência não correlacionada; $a = 2.0$ traço de um movimento browniano fracionário; $a = 2.5$ traço de um movimento browniano fracionário com incrementos de persistência.	103
7.1	Várias séries geradas artificialmente com parâmetros diferenciados, com 2^{20} dados gerados pela equação 6.9. (a) $a = 0,00$, (b) $a = 1,00$, (c) $a = 1,50$, (d) $a = 2,00$, (e) $a = 2,50$, (f) $a = 3,00$	104
7.2	Retornos para as séries investigadas. (a) $a = 0,00$, (b) $a = 1,00$, (c) $a = 1,50$, (d) $a = 2,00$, (e) $a = 2,50$, (f) $a = 3,00$	105
7.3	Densidades de probabilidade (a) $a = 0,00$, (b) $a = 1,00$, (c) $a = 1,50$, (d) $a = 2,00$, (e) $a = 2,50$, (f) $a = 3,00$	106
7.4	Teste da condição de Lindeberg para as séries correlacionadas.	108
7.5	Convergência da <i>skewness</i> para a aplicação do TLC.	108
7.6	Convergência da <i>skewness</i> para a aplicação do TSL com baixa correlação. (a) $a = 0.00$, (b) $a = 0.25$, (c) $a = 0.50$, (d) $a = 0.75$, (e) $a = 1.00$, (f) $a = 1.25$,	109
7.7	Convergência da <i>skewness</i> para a aplicação do TSL com fortes correlações.	109
7.8	Convergência da <i>kurtosis</i> para a aplicação do TLC.	110
7.9	Convergência da <i>kurtosis</i> para a aplicação do TSL com baixa correlação. (a) $a = 0.00$, (b) $a = 0.25$, (c) $a = 0.50$, (d) $a = 0.75$, (e) $a = 1.00$, (f) $a = 1.25$,	111
7.10	Convergência da <i>kurtosis</i> para a aplicação do TSL com fortes correlações.	111
7.11	Teste KS para séries aplicadas ao TLC e com $0 \leq a < 1$	112
7.12	Teste KS para séries aplicadas ao TLC e com $1 \leq a < 2$	112
7.13	Teste KS para séries aplicadas ao TLC e com $2 \leq a \leq 3$	113

LISTA DE FIGURAS

7.14	Teste KS para as séries aplicadas ao TSL com baixa correlação	113
7.15	Teste KS para as séries aplicadas ao TSL com fortes correlações.	114
7.16	Teste KL para séries aplicadas ao TLC e com $a < 1$	114
7.17	Teste KL para séries aplicadas ao TLC e com $1 \leq a < 2$	115
7.18	Teste KL para séries aplicadas ao TLC e com $2 \leq a \leq 3$	115
7.19	Teste KL para as séries aplicadas ao TSL com baixa correlação	116
7.20	Teste KL para as séries aplicadas ao TSL com fortes correlações.	116

LISTA DE TABELAS

2.1	Probabilidades arbitrárias de um exemplo de dado não justo	27
5.1	Valores da <i>skewness</i> e <i>kurtosis</i> para a série original do DEM/USD	90
5.2	Tabela dos valores médios de D_n	94
5.3	Tabela dos valores médios de D_{KL} em cada um dos teoremas, mostrando que mesmo sendo da ordem de 10^{-2} ainda não são nulos	95
7.1	Valores iniciais da <i>skewness</i> para as séries geradas do movimento browniano fracionado	106
7.2	Valores iniciais da <i>kurtosis</i> para as séries geradas do movimento browniano fracionado	107

SUMÁRIO

1	Introdução	21
2	Variáveis Aleatórias e a Estatística de dados	25
2.1	Variáveis Aleatórias	25
2.2	Funções de Probabilidade	26
2.3	Valor médio, desvio padrão e os momentos estatísticos	29
2.3.1	A média	29
2.3.2	A variância e desvio padrão	31
2.3.3	Momentos de mais alta ordem	32
2.3.4	A função característica	35
2.4	Distribuições conjuntas de probabilidade	37
2.5	Algumas Distribuições de Probabilidades	38
2.5.1	Distribuição Binomial	38
2.5.2	Distribuição Gaussiana(Normal)	40
2.5.3	Distribuições α -estáveis de Lévy	47
2.6	Fractalidade em Séries temporais	55
3	O Teorema do Limite Central e as Seções de Lévy	59
3.1	Teorema do Limite Central(TLC)	59
3.1.1	Alguns exemplos de convergências via TLC	62
3.2	As seções de Lévy	65
3.2.1	O Teorema das Seções de Lévy(TSL)	66
3.2.2	TSL aplicado à séries temporais	70
3.3	O Teorema do Limite Central Clássico vs Teorema das Seções de Lévy	71
3.4	Testes de Convergência	71
3.4.1	Momentos estatísticos	72
3.4.2	Expoente de Hürst e a Fractalidade	72
3.4.3	Teste KS	73
3.4.4	Teste KL de divergência	77
4	A Econofísica como exemplo de séries correlacionadas	79
5	O TLC e o TSL aplicadas a séries financeiras	84
5.1	O banco de dados	84
5.2	A convergência	88

SUMÁRIO

6	Traço do movimento Browniano fracionário	100
6.1	Definição do movimento browniano fracionário	100
7	O TLC e o TSL aplicados à séries de sítios de energia correlacionados	104
7.1	Análise da Convergência	107
8	Conclusões e perspectivas	118
A	Formalismo matemático da condição de Lindeberg	120

1

Introdução

Desde os primórdios da civilização, relatos históricos mostram que jogos de azar estão presentes na sociedade tanto como forma de entretenimento, quanto forma de ludibriar alguns tolos para ganhar lucros[1].

Por volta do século XVII, Girolamo Cardano publicou o que se tratava do primeiro livro em que podia-se observar alguma formalidade matemática, que fosse capaz de explicar minimamente as chances por trás dos jogos de azar. Esse livro, intitulado “O Livro dos Jogos de Azar”, apresenta, entre outras coisas as probabilidades existentes no lançar de um dado não viciado, ou mesmo no sacar de uma carta em um baralho normal com 52 cartas[2, 3].

Os trabalhos de Pascal e Fermat, que se correspondiam por cartas[4], anos mais tarde da publicação do livro de Cardano, trouxeram as bases matemáticas que Jakob Bernoulli precisava para escrever seu famoso livro “A arte de Conjecturar” em 1713[5]. Em seu tratado, Bernoulli, além de outras coisas, explora o lançar de uma moeda justa, com uma probabilidade p de, após o lançamento, estar com a face voltada para apenas uma das opções, cara ou coroa.

Nesse estudo, Bernoulli foi capaz de mostrar a lei dos grandes números, onde é afirmado que depois de muitos testes, ou lançamentos, se tem em média p -vezes a face cara, ou coroa, dependendo da escolha prévia para a probabilidade p .

Essa análise pode ser extrapolada para o estudo de situações onde se tem apenas duas escolhas, como cara ou coroa, esquerda ou direita, o que é definido como distribuições binomiais. Por exemplo ao lançar uma moeda quatro vezes é fácil notar que são 16 resultados possíveis, dentre eles, todos os lançamentos serem cara ou todos os lançamentos serem coroa. Utilizando o triângulo de Pascal[6], nota-se que apenas um dos resultados tem apenas cara, 4 resultados tem 3 caras e uma coroa, 6 resultados tem duas caras e duas coroas, 4 resultados tem 3 caras e uma coroa e um resultado tem todas coroas.

Para poucos lançamentos, é simples determinar as probabilidades, mas alguns eventos acabam se tornando inviáveis de calcular diretamente, como por exemplo, ao lançar uma moeda 60 vezes, de quantas formas pode se observar a ocorrência 29 vezes a face cara voltada para cima? Usando uma calculadora a resposta surge como $1,145 \cdot 10^{17}$, algo improvável de ser calculado na época de Bernoulli.

Em meados do século XVIII, Abraham de Moivre publicou o livro chamado, “A doutrina da probabilidade”[7], no qual ele deduzia uma equação aproximada para as probabilidades do lançamento de uma moeda "viciada", ou seja sem a equiparação das chances entre os possíveis resultados, cara ou coroa. Esse trabalho, mais tarde, se tornou a base da função

conhecida como Função Erro, que futuramente se tornaria mais reconhecida como curva normal.

A curva normal, ganhou fama e status, quando sua forma, bem característica de sino, começou a surgir em dados empíricos de trabalhos ligados às ciências sociais e humanas, não apenas em estudos matemáticos e jogos de azar, com os trabalhos do sociólogo Adolphe Quetelet[8], pioneiro em métodos quantitativos em sociologia. Em 1835, ele publicou um livro intitulado, “Sobre o homem e o desenvolvimento de suas faculdades”, no qual foi introduzido o conceito de “homem médio”, alguém que possuía características físicas comuns, médias, sem valores extravagantes. Seu trabalho mostrava a existência de uma distribuição de valores para grandezas quantitativas como altura, peso e etc. Quetelet verificou que essa distribuição não era uniforme em um intervalo de valores, mas sim que, grande parcela da população estava concentrada em um intervalo próximo a um valor médio e os valores fora da média, tinham suas probabilidades decaídas exponencialmente. Quetelet tornou o “homem médio” como uma meta social, afim de que, se a sociedade fosse baseada nessa média, a riqueza também deveria o ser. Riqueza essa que era distribuída por uma lei de potência, como mostrado por Pareto[9], e não uma distribuição similar a mostrada por Quetelet.

Na ciência natural a distribuição normal tornou-se matematicamente formalizada como um comportamento evidente na natureza. No início do século XVII, os trabalhos de Legendre e Gauss trouxeram essa poderosa afirmação. Ambos estavam interessados em estudar os erros comumente recorrentes em análises astronômicas de movimentos de asteroides e cometas que eram muito imprecisas dadas as condições da época e o longo período das órbitas desses corpos[10].

O triunfo de suas análises se deu pelo estudo dos mínimos quadrados[10], um método que consiste em determinar a melhor curva que se ajusta aos dados empíricos abordados. Gauss se tornou bastante famoso ao predizer com muita exatidão, o retorno do asteroide Ceres, feito que era dito como impossível de ser determinado por falta de dados. Ele justificou que os ajustes que ele utilizava no método dos mínimos quadrados, considerava que os dados seguiam uma distribuição normal, que mais a frente, recebeu em homenagem à Gauss, a alcunha de Curva Gaussiana.

Mas porque tantas coisas na natureza seguem a distribuição normal? A resposta para essa pergunta começou a ser respondida ainda no século XVII por Pierre-Simon de Laplace que publicou o texto denominado: “Teoria analítica de probabilidades”, no qual ele mostrava modelos estocásticos e métodos analíticos para suas soluções. Laplace mostrou que em uma distribuição binomial de probabilidades, com variáveis independentes, quando o número de variáveis agregadas tende ao infinito, a distribuição de probabilidades seguia a curva normal.

Seu tratado matemático complexo, envolvia o uso de transformadas de Fourier para determinar a função característica relacionada com a distribuição de probabilidades. Em sua publicação, Laplace extrapola o conceito apresentado pela distribuição binomial e inclusive afirma que: “... a série converge mais rapidamente, quanto mais complicado sua fórmula é, de tal forma que o procedimento se torna mais preciso, quanto mais ele se torna necessário”. Com esse estudo, Laplace traz as bases para o que hoje conhecemos como “Teorema do Limite Central”[11].

A alcunha de “Teorema do Limite Central” foi provavelmente atribuído por Georg Pólya em um artigo de 1920[12], quando o mesmo escreve: “É observado que a distribuição gaussiana surge em diversas situações, e isso se deve a um teorema limitante, que exerce

um papel central na teoria de probabilidade”.

O Teorema do Limite Central é amplamente aplicado nos mais inusitados sistemas, desde velocidades de movimento de partículas em um gás[13] até a altura das pessoas[8]. Mas as variáveis desses sistemas aqui citados, são ditas independentes, como será que se comporta uma série temporal fortemente correlacionada frente a aplicação desse poderoso teorema limitante?

Sistemas complexos[14], de forma geral, tem atraído cada vez mais trabalhos científicos em revistas periódicas, como pode ser observado no portal de periódicos da CAPES, onde os resultados mostram mais de 2 milhões de artigos relacionados com a busca. Tal linha de pesquisa tem sido tão frequentemente abordada por ser rica em propriedades que não são facilmente observadas em outros sistemas, tais como criticalidade auto-organizada, fractalidade e multifractalidade, invariância por escala e algumas outras presentes em vários sistemas biológicos, físicos e até econômicos.

Com todas essas propriedades, os sistemas complexos se tornam ideais para compreender melhor, qual a dinâmica por trás do Teorema do Limite Central e até onde vão as suas afirmações.

Esse caso então é novamente analisado sob uma outra perspectiva, o Teorema das Seções de Lévy. O Teorema das Seções de Lévy, proposto por Paul Lévy[15], trata de uma generalização do conceito abordado no Teorema do Limite Central, onde pode-se mostrar que até mesmo variáveis aleatórias correlacionadas, também seguem um regime de distribuição de probabilidades gaussiano.

Através das ferramentas usualmente utilizadas em técnicas aplicadas em problemas de Mecânica Estatística, que serão melhor abordadas no capítulo 2, pretende-se identificar diversas propriedades que surgem em sistemas econômicos, tais como: multifractalidade, caudas grossas, correlações de longo alcance e não homogeneidade da variância, ou seja, a heterocedasticidade da série, e verificar como as mesmas formam um caráter único nas séries e tornam bastante atrativo o estudo de convergência para o Teorema do Limite Central. Além de séries econômicas, algumas outras séries geradas sob um movimento browniano fracionado, também possuem propriedades interessantes, e uma correlação bem definida que depende de um parâmetro ajustável do modelo, que auxilia na compreensão do mecanismo de convergência.

No capítulo 3 serão abordadas as bases fundamentais que formam o Teorema do Limite central em sua forma clássica, provado por Laplace[11] e a formação proposta por P. Lévy[15], que torna o TLC(Teorema do Limite Central) muito mais geral, podendo ser aplicado em séries fortemente correlacionadas. O teorema generalizado proposto por Lévy separa a série em diversas seções, chamadas de **seções t** , onde o parâmetro regulador da série é a variância da seção t , por isso recebe o nome de Teorema das Seções de Lévy(TSL).

As principais propriedades que emergem da interação dinâmica dos sistemas econômicos serão claramente discutidas no capítulo 4, onde serão apresentadas as tendências acadêmicas para a explicação de tais propriedades. É preciso entender profundamente essas origens que são apresentadas no sistema econômico, para então assim, discutir o que acontece quando os teoremas são aplicados.

No capítulo 5, utilizando séries empíricas da taxa de câmbio do Marco Alemão frente ao dólar americano, serão apresentadas todas as propriedades que essa série temporal apresenta e depois seguir com uma análise de como essas propriedades vão sendo modificadas à medida que os teoremas são aplicados, e como a convergência para a gaussiana se fortalece enquanto que algumas das propriedades características da série temporal são

destruídas, ou mesmo suprimidas, dando lugar ao padrão pertencente a curva normal.

Afim de compreender a dinâmica de convergência, também são analisados alguns outros sistemas com motivação física, que são baseados em um movimento browniano fracionado, onde a correlação dos termos da série é facilmente ajustada e depende de um parâmetro específico. Essas séries serão expostas no capítulo 6 e os resultados sobre a convergência no capítulo 7.

Todos os resultados apresentados nos capítulos 5 e 7, são inéditos na literatura, e parte deles, provenientes das análises iniciais do capítulo 5, o que concerne o estudo da convergência com os parâmetros indicadores de momentos estatísticos, *skewness* e *kurtosis*, da evolução do expoente de Hürst e o espectro multifractal, já foram publicados em uma revista internacional com Qualis A2, segundo o WebQualis da CAPES[16]. Os resultados que competem a análise de convergência através dos testes estatísticos, Kolmogorov-Smirnov(KS) e Kullback-Leibler(KL), e os resultados apresentados no capítulo 7, sobre o movimento browniano fracionado, ainda não foram publicados, ou mesmo expostos em eventos ou congressos.

Esses trabalhos de análise competem um parte importante da construção do conhecimento estatístico de sistemas dinâmicos. Ao compreender melhor a relação que existe entre as variáveis estatísticas e como a convergência ao comportamento gaussiano é obtida em diferentes sistemas e casos, é possível melhorar o tratamento estatístico e computacional utilizado nesses sistemas.

2

Variáveis Aleatórias e a Estatística de dados

Como discutido no capítulo anterior, essa tese está fundamentada no estudo estatístico da convergência de distribuições de probabilidade através da aplicação de dois teoremas fundamentais, o Teorema do Limite Central(TLC) e o Teorema das Secções de Lévy(TSL). Antes desse conteúdo ser devidamente abordado, se faz necessário uma breve visita aos conteúdos básicos de mecânica estatística, afim de que o leitor possa acompanhar a investida sem quaisquer problemas de revisão bibliográfica.

Nesse capítulo serão abordados os princípios que definem variáveis aleatórias, séries temporais, distribuições de probabilidade, assim como, momentos estatísticos, cumulantes e algumas outras propriedades pertinentes a esse estudo.

2.1 Variáveis Aleatórias

No que se concerne ao âmbito da descrição fundamental de variável aleatória, pode-se dizer com confiança que um evento é dito aleatório se é praticamente impossível prevê-lo a partir do seu estado inicial, ou seja, uma variável aleatória tem um valor que está sujeito a chance. É importante compreender que o que é apresentado aqui mostra um caráter bem realista ao conceito de variável aleatória.

Diversos sistemas são característicos de variáveis aleatórias, como por exemplo, o lançar de um dado justo ou uma moeda justa, assim como um sorteio de bingo. Apesar de ditas aleatórias, todas essas variáveis poderiam, em princípio, ser determinadas conhecendo todos os fenômenos ao seu redor, como velocidade do vento, força ao ser lançado, massa do dado, aceleração da gravidade local, até mesmo campos magnéticos locais, seria possível determinar qual face iria ser “sorteada”. Segundo o físico e matemático Laplace, se fosse possível conhecer toda a informação do universo, não existiria o futuro, tudo seria previsto. Como o acesso a todas essas informações por vezes é inviável, essas variáveis são ditas aleatórias, sujeitas ao acaso. Na natureza, é possível encontrar aleatoriedade “verdadeira”, ou “pura”, fenômenos que realmente são sujeitos completamente pelo acaso, são os sistemas quânticos, que fundamentalmente violam o pensamento de Laplace, pois variáveis quânticas não podem ser determinadas previamente, apenas através de medidas.

Uma variável aleatória pode tomar diversos valores possíveis, um dado de 6 lados pode tomar seis valores, uma moeda, dois valores, uma dezena da mega-sena pode assumir 60 valores diferentes. A definição matematicamente rigorosa de uma variável aleatória pode ser descrita como o seguinte conceito: se X é uma variável aleatória em um espaço amostral

E , logo X é uma função que mapeia um elemento do conjunto espaço amostral (E) para um conjunto imagem (I) onde se encontra a sua frequência de ocorrência, sendo assim:

$$X : E \mapsto I \quad (2.1)$$

onde I é um conjunto real de tal maneira que o mapeamento inverso de cada intervalo em I , leva a um evento ocorrido em E . A normalização do conjunto imagem, faz com que os seus elementos assumam valores de tal maneira que a soma de todos eles seja 1, assim, fornecendo a probabilidade de ocorrência da variável.

É possível classificar variáveis aleatórias como discretas e contínuas. As variáveis discretas em sua maioria, são definidas sobre um conjunto finito de ações, ou de valores possíveis a serem sorteados, como novamente por exemplo, o lançar de um dado, onde podemos ter apenas 6 faces possíveis, ou mesmo o valor de um ativo financeiro onde este seu valor é atualizado de forma discreta, em incrementos finitos, com relação a um tempo determinado, gerando uma chamada série temporal discreta, ou seja, onde os valores não são contínuos. Pode-se inferir então que, se a menor diferença não nula entre dois possíveis valores dessa variável for sempre finita, logo ela é uma variável discreta, não está definida em um conjunto contínuo e sim em um arranjo estratificado e finito.

Variáveis aleatórias contínuas, podem assumir valores dentro do conjunto dos reais (\mathbb{R}), não existindo uma sub-divisão finita entre valores consecutivos, ou seja, as variáveis contínuas são aquelas que a menor diferença não nula entre dois possíveis valores da variável aleatória nem sempre toma um valor finito. Como exemplos, é possível citar uma partícula quântica livre, que viaja com uma função de onda como onda-plana por todo o espaço, e sua função é estendida em todo \mathbb{R} , ou um grão de pólen suspenso em água, que move-se conforme o movimento Browniano descrito por Einstein[17] onde sua posição é uma variável aleatória contínua, podendo assumir qualquer valor na reta real.

2.2 Funções de Probabilidade

A cada um dos eventos associados com uma variável aleatória, é possível determinar uma probabilidade associada ao mesmo. Para cada valor de X , a variável aleatória, ou seja, os x_i , existe uma probabilidade p_i de ocorrer. É possível escrever $p_i = f(x_i)$, como a probabilidade que $x = x_i$, e chamar a função $f(x)$, ou mesmo p_i , de função de probabilidade para a variável x .

No caso de variáveis discretas, pode-se dizer que o conjunto de valores, p_i , a distribuição de probabilidade, deve satisfazer as condições $p_i \geq 0$ e de tal forma que seja normalizada:

$$\sum_{i=1}^n p_i = 1,$$

onde n é o número de realizações possíveis da variável, podendo inclusive ser um número muito grande, até mesmo ∞ . É possível definir também uma outra quantidade útil, a função densidade de probabilidade $P_X(x)$, definida como

$$P_X(x) = \sum_{i=1}^n p_i \delta_{x-x_i} \quad (2.2)$$

em que δ_{x-x_i} é a função delta de Kronecker. Um exemplo simples de visualizar a condição é no lançar de um dado não justo, de tal forma que, das 6 faces, as probabilidades sejam dadas pela seguinte tabela:

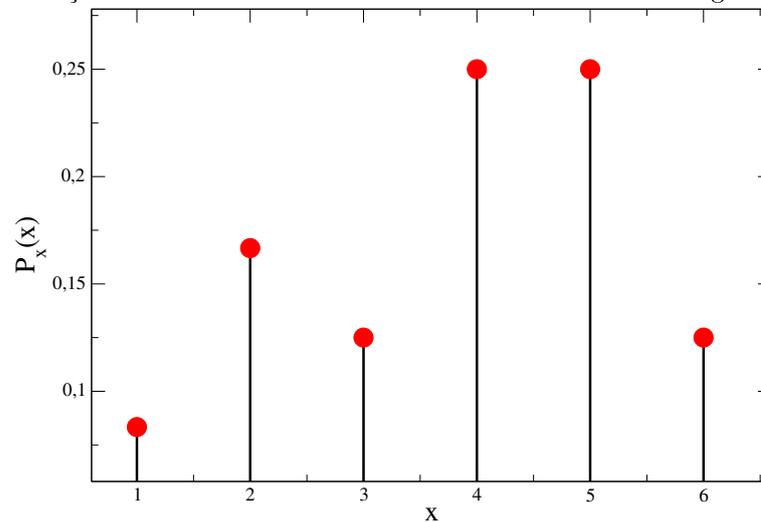
Tabela 2.1: Probabilidades arbitrárias de um exemplo de dado não justo

Face	p_i
1	1/12
2	1/6
3	1/8
4	1/4
5	1/4
6	1/8

Fonte: Autor, 2014.

Dessa forma, a sua densidade de probabilidade seria descrita conforme mostra a figura 2.1

Figura 2.1: Função Densidade de Probabilidade de um dado viciado segundo a tabela 2.1



Fonte: Autor, 2014.

Se ao invés disso, X for uma variável aleatória que pode tomar valores em um conjunto contínuo, assim como um conjunto real, é conhecido que em um intervalo de dois possíveis valores de X , denotados de forma exemplificada, x_1 e x_2 , é possível assumir que existe uma função contínua, $P_X(x)$, de tal forma que a probabilidade de X ter um valor determinado entre x_1 e x_2 é dada pela área sob a curva de $P_X(x)$ versus x , entre $x = a$ e $x = b$, assumindo que, $a < b$, ou seja:

$$Prob(x) = \int_a^b dx P_X(x) \quad (2.3)$$

onde $P_X(x)dx$ é a probabilidade de encontrar X no intervalo entre x e $x+dx$. A densidade de probabilidade contínua, também está sujeita a algumas condições, tais como $P_X(x) \geq 0$ e que

$$\int_{\{E\}} P_X(x)dx = 1,$$

com E sendo o espaço amostral em que a variável se encontra, o que pode de maneira geral significar o intervalo $(-\infty, \infty)$. Um exemplo de distribuição de probabilidade contínua, é o módulo quadrado da função de onda resultante da interação de uma partícula com um potencial harmônico, ou, um oscilador harmônico quântico. Partindo da equação de Schrödinger, independente do tempo, com um potencial dado por

$$V(x) = \frac{1}{2}m\omega^2x^2,$$

temos na base de posições a equação:

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + \frac{1}{2}m\omega^2x^2\psi(x) = E\psi(x) \quad (2.4)$$

e tomando $\alpha = \frac{m\omega}{\hbar}$, a função de onda normalizada é escrita como:

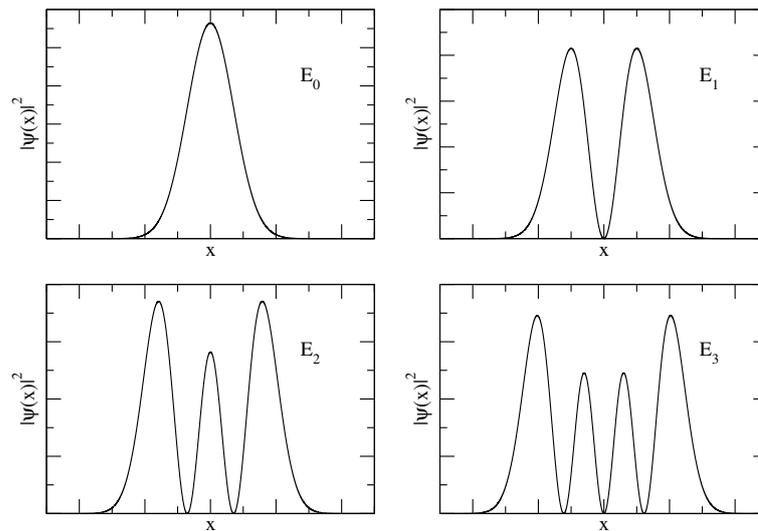
$$\psi_n(x) = \left(\frac{\alpha}{\pi}\right)^{1/4} \frac{1}{\sqrt{2^n n!}} H_n(\sqrt{\alpha}x) e^{-\alpha x^2/2} \quad (2.5)$$

onde $H_n(y)$ são os polinômios de Hermite. Através dessa solução e usando o fato de que a densidade de probabilidade em mecânica quântica é o módulo quadrado da função de onda, ou seja:

$$p(x) = |\psi_n(x)|^2 \quad (2.6)$$

é possível ter uma visão geral dos níveis de energia e da probabilidade de encontrar uma partícula confinada nesse potencial, visualizando a figura 2.2. A probabilidade de encontrar a partícula em uma determinada posição x tem diversas distribuições possíveis, dependendo do nível de energia que se está interessado.

Figura 2.2: Função Densidade de Probabilidade do oscilador harmônico quântico, com parâmetros livres.



Fonte: Autor, 2014.

2.3 Valor médio, desvio padrão e os momentos estatísticos

A função de probabilidade de uma variável aleatória fornece de forma detalhada, diversas informações sobre o comportamento da variável em questão, como valores prováveis, desvios sobre esses pontos, assimetrias na distribuição e outras características, mas nem sempre essa função é facilmente acessível. Para esses casos, é possível tomar um outro tipo de caminho de análise, estudando as propriedades específicas de cada espaço amostral.

2.3.1 A média

A média, comumente confundida com mediana e moda, é a medida da tendência central de uma distribuição, ou seja, o valor em que as variáveis do espaço amostral tendem a seguir. É importante frisar que nem sempre a média é o valor mais provável de ocorrer, esse quantidade é chamada de moda da distribuição. A Mediana, por sua vez, é o valor que divide a distribuição em duas partes iguais. Em algumas distribuições simétricas, esses valores são os mesmos, mas essa é uma exceção, não a regra.

De forma pictórica e didática é possível afirmar que a média indica a posição do centro de massa, o ponto de equilíbrio, de uma densidade de probabilidade, $P_X(X)$. A média pode ser representada de várias maneiras na literatura, podendo ser encontrada como μ , $\langle x \rangle$, \bar{x} ou mesmo $E(x)$, chamado de valor esperado.

Por definição a média aritmética de uma distribuição discreta é a soma de todos os valores obtidos em suas n realizações, dividida pela quantidade total de realizações, ou seja:

$$\mu = \sum_{i=1}^n \frac{x_i}{n} \quad (2.7)$$

sendo assim, pode-se escrever também como:

$$\mu = \sum_{i=1}^n p_i x_i \quad (2.8)$$

Por exemplo, considerando o dado viciado abordado na tabela 2.1, a média seria:

$$\mu_{\text{dado viciado}} = \frac{1}{12} \times 1 + \frac{1}{6} \times 2 + \frac{1}{8} \times 3 + \frac{1}{4} \times 4 + \frac{1}{4} \times 5 + \frac{1}{8} \times 6 = 3,79, \quad (2.9)$$

enquanto que claramente a moda seria 4 ou 5, os dois valores mais prováveis, com 1/4 de chance de ocorrer, e a mediana seria um valor próximo a 4,17, que divide a distribuição em duas partes iguais.

Para distribuições contínuas, o somatório se torna uma integral e portanto, a média pode ser definida como sendo:

$$\mu = \int_{\{E\}} P_X(x) x dx \quad (2.10)$$

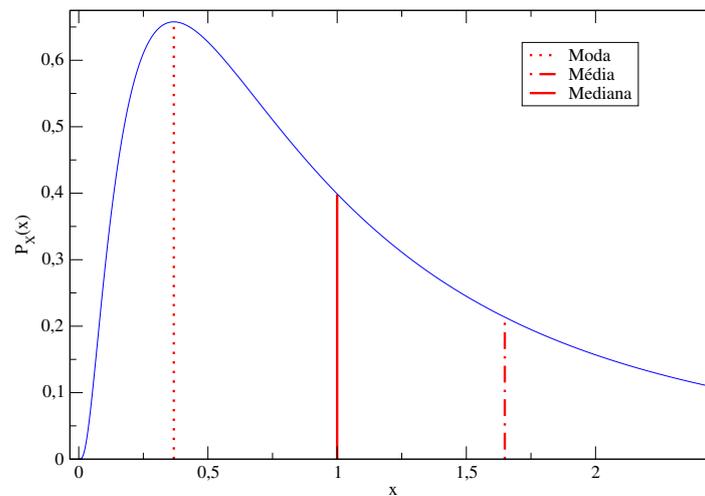
onde $\{E\}$ é todo o espaço amostral, usualmente considerando o intervalo $(-\infty, \infty)$.

Pode-se tomar como exemplo uma distribuição qualquer, como a função:

$$P_X(x) = \frac{1}{x\sqrt{2\pi}} e^{-\frac{[\ln(x)]^2}{2}} \quad (2.11)$$

que tem como média, resolvendo a integral da equação 2.10, $e^{1/2}$, mediana sendo 1, obtida integrando a função de $-\infty$ até a afim obter probabilidade de 0,5, e a moda, calculada encontrando o ponto onde a derivada é nula, é dada por e^{-1} . É possível ver a identificação de cada um desses pontos na figura 2.3.

Figura 2.3: Função Densidade de Probabilidade da função descrita na equação 2.11 onde são indicados os valores da média, mediana e moda, podendo identificar as características de cada uma delas, especialmente a moda onde claramente é o ponto mais alto da probabilidade. A mediana divide a distribuição em duas partes iguais, ou seja, a soma das probabilidades de os valores possíveis da variável de $-\infty$ até 1, é igual a soma de todas as probabilidades dos valores acima de 1 até ∞ . E a média, calculada pela equação 2.10, é a tendência, ou centro de massa, da distribuição.



Fonte: Autor, 2014.

2.3.2 A variância e desvio padrão

O desvio padrão por sua vez, é definido a partir de outra quantidade, a variância. Ele define a dispersão que existe nos dados, de forma grosseira, a largura da distribuição, ou então, o quão distantes os valores podem ser encontrados fora da média, ou seja, fora da tendência. Esses desvios tanto podem ser positivos quanto negativos, ao invés de calcular apenas a diferença entre eles, a definição de variância é calculada com o quadrado das diferenças, afim de evitar uma variância nula sempre, sendo assim, variância por definição será:

$$\text{Var}(x) = \sum_{i=1}^n (x_i - \mu)^2 p_i \quad (2.12)$$

Definida a variância é possível agora determinar o desvio padrão, que indica o quanto as medidas estão dispersas. Essa grandeza estatística é definida como sendo a raiz quadrada da variância, ou seja:

$$\sigma_x = \sqrt{\text{Var}(x)} = \sqrt{\sum_{i=1}^n (x_i - \mu)^2 p_i} \quad (2.13)$$

Usando mais uma vez o exemplo do dado tendencioso, é possível determinar qual valor da variância e do desvio padrão:

$$\begin{aligned} \text{Var}_{\text{dado viciado}} &= (1 - 3,79)^2 \times \frac{1}{12} + (2 - 3,79)^2 \times \frac{1}{6} + (3 - 3,79)^2 \times \frac{1}{8} \\ &+ (4 - 3,79)^2 \times \frac{1}{4} + (5 - 3,79)^2 \times \frac{1}{4} + (6 - 3,79)^2 \times \frac{1}{8} \\ &= 2,25 \end{aligned} \quad (2.14)$$

$$\sigma_{\text{dado viciado}} = \sqrt{2,25} = 1,5 \quad (2.15)$$

É possível observar que os valores possíveis estão bem dispersos em relação a média, com um desvio padrão de 1,5, ou seja, a tendência do sistema é a média de 3,79, com um desvio de 1,5.

No caso contínuo, novamente o somatório é agora dado por uma integral em todo o espaço, de tal forma que a variância e o desvio padrão são definidos como:

$$\text{Var}(x) = \int_{\{E\}} P_X(x)(x - \mu)^2 dx \quad (2.16)$$

$$\sigma_x = \sqrt{\text{Var}(x)} = \sqrt{\int_{\{E\}} P_X(x)(x - \mu)^2 dx} \quad (2.17)$$

Resolvendo essa integral, equação 2.16, para a a função particular dada pela equação 2.11, apresentada como exemplo anterior de uma distribuição contínua, o resultado fornece uma variância e desvio padrão de:

$$\text{Var}(x) = e(e - 1) \quad (2.18)$$

$$\sigma_x = \sqrt{e(e - 1)} \quad (2.19)$$

que são os valores dos desvios obtidos quando é utilizada essa distribuição de probabilidade.

2.3.3 Momentos de mais alta ordem

Além das características abordadas na média, as distribuições de probabilidade possuem diversas outras peculiaridades, como a forma e contorno da distribuição, que dependem de algumas grandezas. De forma generalizada, as quantidades que especificam as nuances de cada distribuição, são chamados momentos estatísticos e de maneira generalizada são descritos da seguinte forma:

$$\mu_m = \sum_{i=1}^n x_i^m p_i \quad (2.20)$$

$$\mu_m = \int_{\{E\}} P_X(x) x^m dx \quad (2.21)$$

onde a equação 2.20 se refere a distribuições discretas e a equação 2.21 às distribuições contínuas. Como informado, cada momento, indicado por um expoente específico nas equações que representa também a ordem do momento, indica um parâmetro da curva, desde sua largura, à diversos graus de assimetria e mesmo o formato que a distribuição tem. Os momentos pares, estão relacionados com a forma que a distribuição tem e os momentos ímpares são relacionados com as assimetrias que a distribuição tem acerca da média.

Momentos centrados na média

Uma outra definição para os momentos, são os momentos centrados na média. Esses momentos são úteis para calcular a variância, *skewness* e *kurtosis*, que são quantidades relevantes na análise quantitativa do formato específico de cada distribuição. Os momentos centrados na média são definidos como:

$$\mu'_m = \langle (x - \langle x \rangle)^m \rangle \quad (2.22)$$

onde por definição $\langle f(x) \rangle$ é o valor esperado da função, ou seja, para funções discretas, dado pela equação:

$$\langle f(x) \rangle = \sum_{i=1}^n p_i f(x_i), \quad (2.23)$$

e para funções contínuas é dado por

$$\langle f(x) \rangle = \int_{\{E\}} P_X(x) f(x) dx, \quad (2.24)$$

O primeiro momento centrado na média, obviamente é zero, pois se trataria da subtração entre a média e ela mesma. Para o segundo momento centrado na média o resultado é a variância, pois, como já era de se esperar, é a dispersão em primeira ordem dos dados, e por si só, ao fazer $m = 2$, a equação 2.22 se torna exatamente igual a equação 2.12, para o caso discreto. É possível mostrar a partir da definição fornecida na equação 2.22 que a própria variância depende do segundo momento calculado pelas equações 2.20 ou 2.21.

$$\begin{aligned} \mu'_2 &= \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 - 2x\langle x \rangle + \langle x \rangle^2 \rangle \\ &= \langle x^2 \rangle - 2\langle x \rangle \langle x \rangle + \langle x \rangle^2 \\ &= \langle x^2 \rangle - 2\langle x \rangle^2 + \langle x \rangle^2 \\ &= \langle x^2 \rangle - \langle x \rangle^2 \end{aligned} \quad (2.25)$$

Tomando o fato de que $\langle \langle x \rangle \rangle = \langle x \rangle$, ou seja, o valor esperado do valor esperado da variável x , é o seu próprio valor esperado.

Como dito, se faz interessante mostrar as relações existentes para alguns outros momentos de ordem mais altas, como o terceiro e o quarto momentos, que estão relacionados com a assimetria da distribuição e o seu grau de abaulamento. Usando a definição fornecida na equação 2.22, é possível verificar que:

$$\begin{aligned}\mu'_3 &= \langle (x - \langle x \rangle)^3 \rangle = \langle x^3 - 3x^2\langle x \rangle + 3x\langle x \rangle^2 - \langle x^3 \rangle \rangle \\ &= \langle x^3 \rangle - 3\langle x^2 \rangle \langle x \rangle + 3\langle x \rangle \langle x \rangle^2 - \langle x^3 \rangle \\ &= \langle x^3 \rangle - 3\langle x^2 \rangle \langle x \rangle + 2\langle x \rangle^3\end{aligned}\quad (2.26)$$

$$\begin{aligned}\mu'_4 &= \langle (x - \langle x \rangle)^4 \rangle \\ &= \langle x^4 - 4x^3\langle x \rangle + 6x^2\langle x \rangle^2 - 4x\langle x \rangle^3 + \langle x \rangle^4 \rangle \\ &= \langle x^4 \rangle - 4\langle x^3 \rangle \langle x \rangle + 6\langle x^2 \rangle \langle x \rangle^2 - 4\langle x \rangle \langle x \rangle^3 + \langle x \rangle^4 \\ &= \langle x^4 \rangle - 4\langle x^3 \rangle \langle x \rangle + 6\langle x^2 \rangle \langle x \rangle^2 - 3\langle x \rangle^4\end{aligned}\quad (2.27)$$

Esses são alguns dos momentos centrados na média. Com esses resultados, são definidas as grandezas *skewness* e *kurtosis* que são momentos padrões, definidos segundo:

$$\gamma_m = \frac{\mu'_m}{\sigma_x^m} \quad (2.28)$$

ou seja, o momento centrado na média, dividido pelo desvio padrão elevado ao grau do momento. A vantagem de definir essas quantidades dessa maneira específica é garantir que: $\gamma_2 = 1$, fornecendo uma outra normalização os momentos e facilitando cálculos posteriores.

Skewness

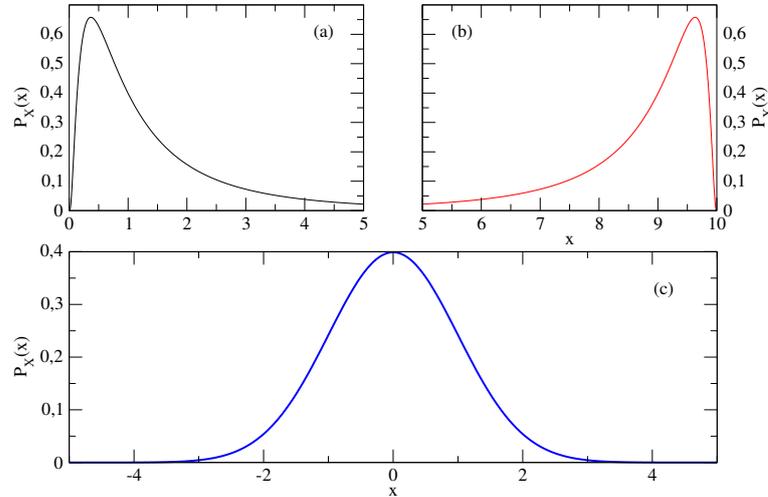
A *skewness*, tem origem na palavra inglesa *skew*, que significa oblíquo ou enviesado. Ao tratar *skewness*, temos a forma oblíqua, ou assimétrica de uma distribuição, ou seja, a *skewness* é a medida da assimetria de uma distribuição de probabilidade sobre a sua média. Seu valor usualmente é positivo, indicando uma cauda a direita da média, ou seja a densidade de probabilidade é mais alongada ou até mesmo mais “gorda” nesse sentido, o valor pode ser negativo, indicando o contrário, ou zero, com um grau de assimetria nulo. Algumas distribuições não possuem uma *skewness* bem definida, por sua própria natureza, por não apresentar assimetria facilmente medida.

Pela equação 2.28 a *skewness* é definida como:

$$s = \frac{\mu'_3}{\sigma_x^3}, \quad (2.29)$$

usando o terceiro momento pois o mesmo é nulo para distribuições simétricas, positivo para distribuições assimétricas à direita e negativo para distribuições assimétricas à esquerda.

Figura 2.4: Nessa figura são apresentadas algumas distribuições que possuem variados valores de *skewness*, podendo mostrar suas propriedades assimétricas. (a) Distribuição com *skewness* positiva, segundo a equação 2.11, mostrando uma assimetria à direita. (b) Distribuição com *skewness* negativa, do espelho da função dada pela equação 2.11 com $x = -x$, mostrando uma assimetria à esquerda. (c) Distribuição com *skewness* nula, da distribuição normal, completamente simétrica sobre o eixo da média.



Fonte: Autor, 2014.

Kurtosis

A *kurtosis* tem origem na palavra grega $\kappa\upsilon\rho\tau\varsigma$ ou, *kyrtos*, que significa curvada, ou arqueada. O objetivo em se calcular a *kurtosis* é obter um parâmetro que indique a forma da distribuição, de maneira grosseira e coloquial, ela representa o grau de afinamento, ou arqueamento da distribuição, medindo se é uma distribuição mais curvada em seu pico, ou se é mais afilada.

Assim como no caso anterior, existem distribuições com *kurtosis* positivas, as chamadas leptocúrticas, mais afuniladas no pico, as que possuem *kurtosis* negativa, chamadas platicúrticas, com o pico mais abaulado, e as distribuições chamadas mesocúrticas, com a *kurtosis* nula.

Definida pela equação 2.28, a *kurtosis* pode ser escrita como:

$$\kappa = \frac{\mu'_4}{\sigma_x^4} \quad (2.30)$$

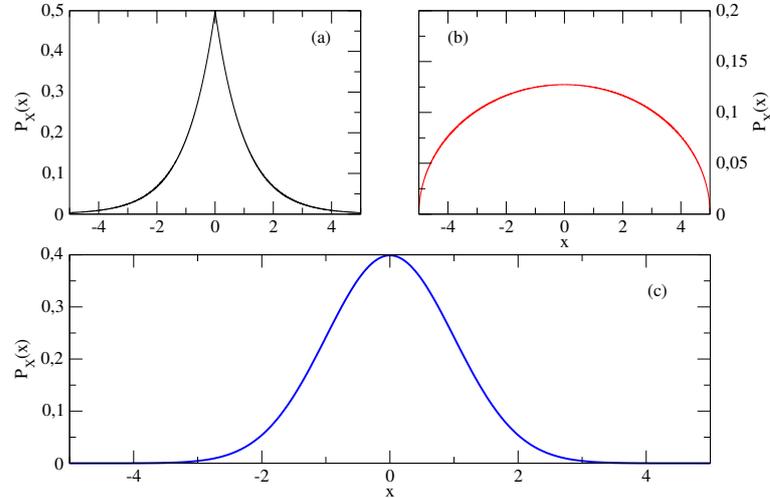
A distribuição normal, ou distribuição gaussiana, é um dos principais motivadores ao uso do quarto momento, pois, possui um valor bem definido de *kurtosis*. Tomando esse valor como base é definida a *kurtosis* excessiva, afim de que se tenha um parâmetro nulo de comparação. A *kurtosis* excessiva é definida por:

$$\begin{aligned} \kappa'_{excessiva} &= \frac{\mu'_4}{\sigma_x^4} - \kappa_{gaussiana} \\ &= \frac{\mu'_4}{\sigma_x^4} - 3 \end{aligned} \quad (2.31)$$

É possível observar na figura 2.5 alguns exemplos de distribuições com vários valores de

kurtosis , para as distribuições de Laplace¹, Wigner² e Gauss.

Figura 2.5: Diferentes distribuições com variados valores de *kurtosis*. (a) Distribuição com *kurtosis* excessiva positiva, segundo a distribuição de Laplace, com *kurtosis* excessiva igual a 3. (b) Distribuição com *kurtosis* excessiva negativa, segundo a distribuição de Wigner, com *kurtosis* excessiva igual a -1. (c) Distribuição com *kurtosis* excessiva nula, a distribuição gaussiana.



Fonte: Autor, 2014.

2.3.4 A função característica

A função característica é uma outra maneira de acessar as propriedades da distribuição de probabilidades. A função característica é definida a partir da transformada de Fourier da própria densidade de probabilidade em questão. Sendo assim, por sua própria definição, a função característica é dada por:

$$f_X(k) = \int_{\{E\}} P_X(x) e^{ikx} dx \quad (2.32)$$

Uma das principais vantagens de estar no espaço de momentos é a expansão em série de Taylor, no caso especial centrada em zero (a chamada série de Maclaurin), para a exponencial. A série de Taylor, tem por definição a forma:

$$f(x) = f(a) + \frac{df(a)}{dx} \frac{(x-a)}{1!} + \frac{d^2f(a)}{dx^2} \frac{(x-a)^2}{2!} + \frac{d^3f(a)}{d(x-a)^3} \frac{x^3}{3!} + \dots, \quad (2.33)$$

para a exponencial, tipo e^y , é fácil notar que o resultado será:

$$e^y = 1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + \dots = \sum_{\alpha=0}^{\infty} \frac{y^\alpha}{\alpha!} \quad (2.34)$$

sendo assim, para o caso da equação 2.32, a expansão se dá por:

$$f_X(k) = \int_{\{E\}} P_X(x) \left(1 + ikx + \frac{(ikx)^2}{2!} + \frac{(ikx)^3}{3!} + \dots \right) dx \quad (2.35)$$

¹Com o caso específico da equação $P_X(x) = \frac{1}{2}e^{-|x|}$

²Com o caso específico da equação $P_X(x) = \frac{2}{25\pi} \sqrt{25 - x^2}$

e aplicando a propriedade distributiva, é fácil ver que:

$$f_X(k) = \int_{\{E\}} P_X(x) dx + ik \int_{\{E\}} P_X(x) x dx + \frac{(ik)^2}{2!} \int_{\{E\}} P_X(x) x^2 dx + \dots \quad (2.36)$$

e por definição,

$$\begin{aligned} f_X(k) &= 1 + ik \langle x \rangle + \frac{(ik)^2}{2!} \langle x^2 \rangle + \frac{(ik)^3}{3!} \langle x^3 \rangle + \dots \\ &= \sum_{m=0}^{\infty} \frac{(ik)^m \langle x^m \rangle}{m!} \end{aligned} \quad (2.37)$$

e portanto, a função característica pode ser obtida utilizando os momentos e por sua vez, usando uma transformada de Fourier, é possível obter a distribuição original. Além dos momentos estatísticos, a função característica é capaz de fornecer diretamente os momentos centrados na média, através da expansão em cumulantes. Para tal, se faz a expansão em série de Maclaurin da função: $\ln f_X(k)$, de tal forma que, as derivadas são dadas por:

$$\left. \frac{d \ln f_X(k)}{dk} \right|_{k=0} = \left. \frac{f'_X(k)}{f_X(k)} \right|_{k=0} = \frac{f'_X(0)}{f_X(0)} = \frac{\langle x \rangle}{-i} = i \langle x \rangle \quad (2.38)$$

$$\begin{aligned} \left. \frac{d^2 \ln f_X(k)}{dk^2} \right|_{k=0} &= \left. \frac{f''_X(k) f_X(k) - (f'_X(k))^2}{f_X^2(k)} \right|_{k=0} = -\langle x^2 \rangle + \langle x \rangle^2 \\ &= i^2 (\langle x^2 \rangle - \langle x \rangle^2) \end{aligned} \quad (2.39)$$

$$\begin{aligned} \left. \frac{d^3 \ln f_X(k)}{dk^3} \right|_{k=0} &= \left. \frac{f'''_X(k) f_X^3(k) - 3f''_X(k) f'_X(k) f_X(k) - 2(f'_X(k))^3 f_X(k)}{f_X^4(k)} \right|_{k=0} \\ &= -i \langle x^3 \rangle + 3i \langle x^2 \rangle \langle x \rangle - 2i \langle x \rangle^3 \\ &= i^3 (\langle x^3 \rangle - 3 \langle x^2 \rangle \langle x \rangle + 2 \langle x \rangle^3) \end{aligned} \quad (2.40)$$

$$\begin{aligned} \left. \frac{d^4 \ln f_X(k)}{dk^4} \right|_{k=0} &= \left. \frac{f_X^{IV}(k) f_X(k) - f_X'''(k) f'_X(k)}{f_X^2(k)} \right|_{k=0} - \\ &- 3 \left(\left. \frac{f_X'''(k) f'_X(k) f_X^2(k) + (f_X''(k))^2 f_X^2(k) - 2f_X''(k) (f'_X(k))^2 f_X(k)}{f_X^4(k)} \right) \right|_{k=0} + \\ &+ 6 \left(\left. \frac{f_X''(k) (f'_X(k))^2 f_X^3(k) - f_X^2(k) (f'_X(k))^4}{f_X^6(k)} \right) \right|_{k=0} \\ &= \langle x^4 \rangle - 4 \langle x^3 \rangle \langle x \rangle - 3 \langle x^2 \rangle^2 + 12 \langle x \rangle^2 \langle x^2 \rangle - 6 \langle x \rangle^4 \\ &= i^4 (\langle x^4 \rangle - 4 \langle x^3 \rangle \langle x \rangle - 3 \langle x^2 \rangle^2 + 12 \langle x \rangle^2 \langle x^2 \rangle - 6 \langle x \rangle^4) \end{aligned} \quad (2.41)$$

e dessa forma, é possível escrever que:

$$\ln f_X(k) = \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} C_m(x) \quad (2.42)$$

onde $C_m(x)$ são os cumulantes, que podem ser trivialmente relacionados com os momentos

centrados na média:

$$C_1(X) = \langle x \rangle \quad (2.43)$$

$$C_2(X) = \langle x^2 \rangle - \langle x \rangle^2 = \langle (x - \langle x \rangle)^2 \rangle = \mu'_2 \quad (2.44)$$

$$C_3(X) = \langle x^3 \rangle - 3\langle x \rangle \langle x^2 \rangle + 2\langle x \rangle^3 = \langle (x - \langle x \rangle)^3 \rangle = \mu'_3 \quad (2.45)$$

$$C_4(X) = \langle x^4 \rangle - 3\langle x^2 \rangle^2 - 4\langle x \rangle \langle x^3 \rangle + 12\langle x^2 \rangle \langle x \rangle^2 - 6\langle x \rangle^4 \quad (2.46)$$

$$= \langle (x - \langle x \rangle)^4 \rangle - 3(\langle x^2 \rangle - \langle x \rangle^2) = \mu'_4 - 3\mu'_2 \quad (2.47)$$

⋮

$$(2.48)$$

Utilizando essa forma é possível ter acesso a diversas informações acerca da distribuição, mesmo que a função $P_X(x)$ seja muito difícil de ser acessada, como é o caso comum em séries temporais, conjuntos de variáveis aleatórias que são sorteadas com o passar do tempo.

2.4 Distribuições conjuntas de probabilidade

Quando duas ou mais variáveis aleatórias estão definidas no mesmo espaço de probabilidade, no mesmo espaço amostral, essas variáveis podem ser descritas através de uma única distribuição que é uma função de todas as variáveis aleatórias. Essa função é construída ao observar as características existentes na ocorrência dessas variáveis e na dependência que existe entre as realizações das diferentes variáveis.

Sem perda de generalidade é trivial entender que se duas variáveis aleatórias, X e Y são distribuídas conjuntamente, ou seja, sobre o mesmo espaço de probabilidade, a distribuição conjunta de probabilidade deverá satisfazer:

$$0 \leq P_{XY}(x, y) \leq 1 \quad (2.49)$$

$$\int_{\{E\}} \int_{\{E\}} P_{XY}(x, y) dx dy = 1 \quad (2.50)$$

e também é possível reduzir a distribuição, para apenas $P_X(x)$, quando as variáveis são independentes, aplicando:

$$P_X(x) = \int_{\{E\}} P_{XY}(x, y) dy \quad (2.51)$$

possibilitando também de maneira similar, determinar $P_Y(y)$. Com essa mesma metodologia, é possível também encontrar os momentos estatísticos associados a variável X :

$$\langle x^m \rangle = \int_{\{E\}} \int_{\{E\}} P_{XY}(x, y) x^m dx dy \quad (2.52)$$

assim como é possível encontrar o momento conjunto, definido como:

$$\langle x^m y^\beta \rangle = \int_{\{E\}} \int_{\{E\}} P_{XY}(x, y) x^m y^\beta dx dy \quad (2.53)$$

Através dessa definição, dois momentos conjuntos podem ser descritos, a covariância e a correlação. A covariância, é uma medida de como as duas variáveis se afastam do

primeiro momento conjunto, e por isso, se mostra muito parecida com a variância, definida pela equação 2.12, de tal forma que:

$$\text{Cov}(X, Y) = \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle = \langle xy \rangle - \langle x \rangle \langle y \rangle \quad (2.54)$$

A correlação, é a medida do grau de relação que essas variáveis tem, o quanto uma depende do resultado da outra, como elas se influenciam. A função correlação está limitada no intervalo $[-1, 1]$, indicando que o valor máximo é a dependência total e uma chamada persistência em repetir eventos e o valor mínimo uma anti-persistência, e se caso as duas variáveis são ditas independentes, logo $\text{Cor}(X, Y) = 0$. Ela é definida a partir da seguinte equação:

$$\text{Cor}(X, Y) = \frac{\langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle}{\sigma_x \sigma_y}. \quad (2.55)$$

onde σ_x e σ_y são os desvios padrões das variáveis aleatórias X e Y .

Para variáveis distribuídas conjuntamente, também é possível determinar uma função característica, com uma definição que surge como uma generalização da equação 2.32:

$$f_Z(k) = \int_{\{E\}} \int_{\{E\}} P_{XY}(x, y) e^{ikG(x,y)} dx dy \quad (2.56)$$

onde a nova variável aleatória Z , depende de X e Y , segundo a função: $z = G(x, y)$.

2.5 Algumas Distribuições de Probabilidades

Para diversos sistemas, como lançar de dados não viciados, colisões de núcleos atômicos, velocidade de partículas em um gás, ou mesmo o simples lançar de uma moeda por diversas vezes, a distribuição de probabilidade $P_X(x)$ é bem conhecida e faz parte de um conjunto de distribuições formalmente representadas com diversas propriedades características. Na presente seção, serão apresentadas apenas algumas distribuições, àquelas que didaticamente são mais compatíveis com o ensaio apresentado na tese.

Inicialmente será abordada a distribuição binomial, de cunho educacional formidável por apresentar um formalismo matemático de compreensão mais simplificada, mas, de contra partida, com aplicações limitadas à sistemas de dois níveis não-correlacionados. Em seguida como uma extrapolação do modelo binomial, será apresentada a distribuição gaussiana, ou curva normal que é um dos principais enfoques do trabalho como muito discutido no capítulo anterior. No fim da seção é formalizada uma generalização das distribuições gaussianas, propostas por Paul Lévy, as distribuições α -estáveis de Lévy[15].

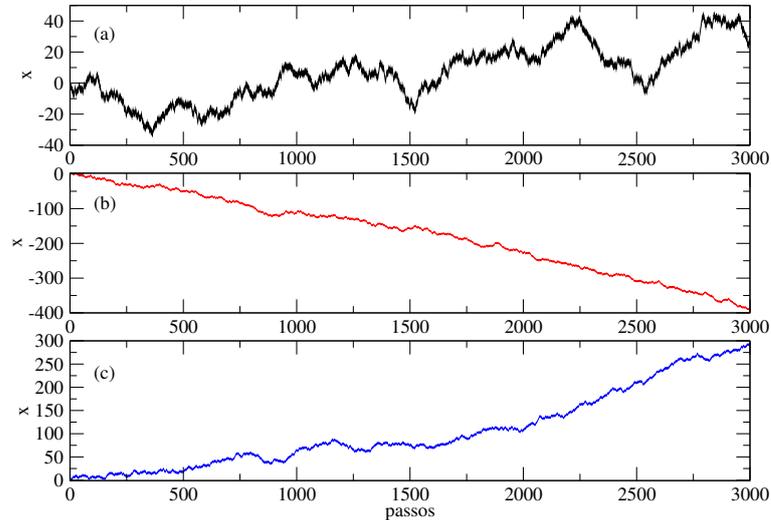
2.5.1 Distribuição Binomial

Em um espaço amostral composto apenas dois possíveis eventos, como o lançar de uma moeda, um lado com probabilidade p e o outro lado com com probabilidade q , ou seja, $1 - p$, surge uma distribuição binomial para um finito número de realizações. Como dito por exemplo o lançar de uma moeda justa, onde $p = q = 0.5$, de forma normalizada.

Um outro exemplo bastante comum, caso esse conjunto amostral seja composto por um móvel, ou caminhante, que é permitido dar passos de tamanho fixo l , tanto para a direita do eixo, valores positivos, quanto para esquerda do eixo, valores negativo, ou seja, restrito a caminhar unidimensionalmente, com probabilidades descritas como no parágrafo acima,

define-se n_+ como o número de passos para o sentido positivo do eixo e n_- o número de passos que ele dá para a direção negativa do eixo, de tal forma que p é a probabilidade dele dar um passo na direção positiva do eixo. A figura 2.6 mostra uma série de caminhantes como esse, com diversos valores de p .

Figura 2.6: Posição x de caminhantes aleatórios com passos de tamanho $l = 1$ e com probabilidades: (a) 0.5, (b) 0.45 e (c) 0.55, em função do número de passos dados.



Fonte: Autor, 2014.

Dessa forma, a probabilidade que ele “caminhe” n_+ passos para a direção positiva do eixo após N passos dados é a combinação de todas as possibilidades multiplicada pela probabilidade de cada evento ocorrer:

$$W_N(n_+) = \frac{N!}{n_+!n_-!} p^{n_+} q^{n_-}. \quad (2.57)$$

Onde o primeiro termo é composto por todas as combinações possíveis com n_+ e n_- , através de combinações simples.

Como $N = n_+ + n_-$, a equação pode ser reescrita de forma a reduzir a quantidade de variáveis:

$$W_N(n_+) = \frac{N!}{n_+!(N - n_+)!} p^{n_+} q^{N - n_+}. \quad (2.58)$$

Que nada mais é que o $(n_+ + 1)$ -ésimo termo da expansão do binômio $(q + p)^N$.

Definindo n , uma posição qualquer no espaço discreto, como sendo:

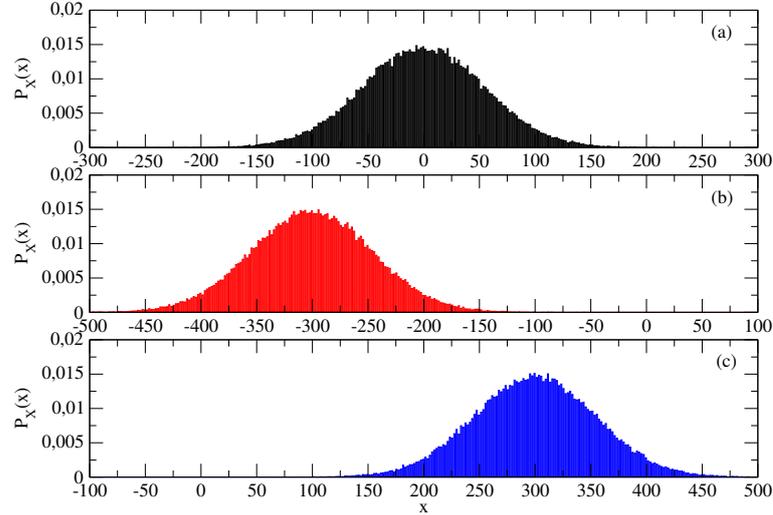
$$n_+ = \frac{N + n}{2} \quad n_- = \frac{N - n}{2}, \quad (2.59)$$

pela equação 2.57, a probabilidade de encontrar o caminhante em uma posição n qualquer do espaço é dada por:

$$P(n) = \frac{N!}{[(N + n)/2]![(N - n)/2]!} p^{(N+n)/2} q^{(N-n)/2}. \quad (2.60)$$

que é uma das formas mais encontradas da distribuição binomial na literatura, onde é possível definir com precisão a probabilidade de encontrar o caminhante na posição n . Na figura 2.7 são mostradas as distribuições de probabilidades do exemplo fornecido na figura 2.6.

Figura 2.7: Probabilidade $P_X(x)$ de caminhantes aleatórios com passos de tamanho $l = 1$ e com probabilidades p : (a) 0.5, (b) 0.45 e (c) 0.55, versus a posição x



Fonte: Autor, 2014.

Para um número de passos N , muito grande, pode-se fazer uma expansão do termo fatorial usando a aproximação de Stirling[18]. Essa expansão é melhor tratada na próxima seção, onde é apresentada a distribuição gaussiana.

2.5.2 Distribuição Gaussiana(Normal)

A distribuição normal pode ser demonstrada como um caso da distribuição binomial quando existe um grande número de realizações da variável aleatória, ou passos na caminhada. Para tal demonstração é importante investigar melhor a probabilidade $W(n_+)$, que é descrita pela equação 2.58. Nesse limite, com $N \rightarrow \infty$, determina-se a expansão em série da função em torno de um valor máximo, com $n_+ = \tilde{n}_+$. Para isso, pode-se fazer uma expansão em série de Taylor, equação 2.33, em torno desse ponto, para $\ln W_N(n_+)$.

Tomando o logaritmo natural da probabilidade,

$$\ln W_N(n_+) = \ln N! - \ln n_+! - \ln(N - n_+)! + n_+ \ln p + (N - n_+) \ln q$$

e utilizando a fórmula de Stirling, $\ln K! \simeq K \ln K - K$, a equação ficará:

$$\begin{aligned} \ln W_N(n_+) \simeq & N \ln N - N - n_+ \ln n_+ + n_+ - (N - n_+) \ln(N - n_+) + \\ & +(N - n_+) + n_+ \ln p + (N - n_+) \ln q \end{aligned} \quad (2.61)$$

Tomando a aproximação como identidade e assim ignorando termos de mais alta ordem da expansão(fórmula de Stirling), a primeira derivada é dada por:

$$\frac{\partial}{\partial n_+} \ln W_N(n_+) = -\ln n_+ + \ln(N - n_+) + \ln p - \ln q. \quad (2.62)$$

Considerando que em \tilde{n}_+ exista um máximo local da função, logo pela condição de máximo da função: $\left. \frac{\partial}{\partial n_+} \ln W_N(n_+) \right|_{n_+=\tilde{n}_+} = 0$, e portanto

$$\ln \frac{\tilde{n}_+}{N - \tilde{n}_+} = \ln \frac{p}{q} \Rightarrow \tilde{n}_+ = \frac{Np}{q} - \frac{\tilde{n}_+ p}{q} \Rightarrow \tilde{n}_+ = \frac{Np}{p+q} = Np \quad (2.63)$$

Calculando agora a segunda derivada para a expansão, para $n_+ = \tilde{n}_+$:

$$\begin{aligned} \left. \frac{\partial^2}{\partial n_+^2} \ln W_N(n_+) \right|_{n_+=\tilde{n}_+} &= \frac{\partial}{\partial n_+} \left(\left. \frac{\partial}{\partial n_+} \ln W_N(n_+) \right|_{n_+=\tilde{n}_+} \right) \\ &= -\frac{1}{Np} - \frac{1}{N - Np} = -\frac{1}{Npq} \end{aligned} \quad (2.64)$$

Desprezando os termos de mais altas derivadas, pois N^m estará presente na $(m - 1)$ -ésima derivada, fazendo com que seu valor seja próximo a nulidade, a expansão em série de Taylor em torno de \tilde{n}_+ , toma a forma:

$$\begin{aligned} \ln W_N(n_+) &= \ln W_N(\tilde{n}_+) + (n_+ - \tilde{n}_+) \left. \frac{\partial \ln W_N(n_+)}{\partial n_+} \right|_{n_+=\tilde{n}_+} + \\ &\quad \frac{(n_+ - \tilde{n}_+)^2}{2!} \left. \frac{\partial^2 \ln W_N(n_+)}{\partial n_+^2} \right|_{n_+=\tilde{n}_+} \\ &= \ln W_N(\tilde{n}_+) - \frac{(n_+ - Np)^2}{2Npq} \end{aligned} \quad (2.65)$$

com o fato de que $\left. \frac{\partial \ln W_N(n_+)}{\partial n_+} \right|_{n_+=\tilde{n}_+} = 0$, agrupando os termos em logaritmo ao lado esquerdo da igualdade:

$$\begin{aligned} \ln \frac{W_N(n_+)}{W_N(\tilde{n}_+)} &= -\frac{(n_+ - Np)^2}{2Npq} \\ W_N(n_+) &= W_N(\tilde{n}_+) e^{-\frac{(n_+ - Np)^2}{2Npq}} \end{aligned} \quad (2.66)$$

Ainda com $W_N(\tilde{n}_+)$ desconhecido. Utilizando da condição de normalização que afirma que a integral da densidade probabilidade em todo espaço deve ser igual a unidade, logo

$$\int_{-\infty}^{\infty} W_N(\tilde{n}_+) e^{-\frac{(n_+ - Np)^2}{2Npq}} dn_+ = 1 \Rightarrow W_N(\tilde{n}_+) = \frac{1}{\int_{-\infty}^{\infty} e^{-\frac{(n_+ - Np)^2}{2Npq}} dn_+} = \frac{1}{\sqrt{2\pi Npq}}. \quad (2.67)$$

Sendo assim, a probabilidade para N tendendo ao infinito, será a chamada distribuição gaussiana, dada por:

$$W_{N \rightarrow \infty}(n_+) = P_G(n_+) = \frac{1}{\sqrt{2\pi Npq}} e^{-\frac{(n_+ - Np)^2}{2Npq}}. \quad (2.68)$$

De forma mais geral, para uma variável aleatória X , a distribuição de probabilidades é dada por:

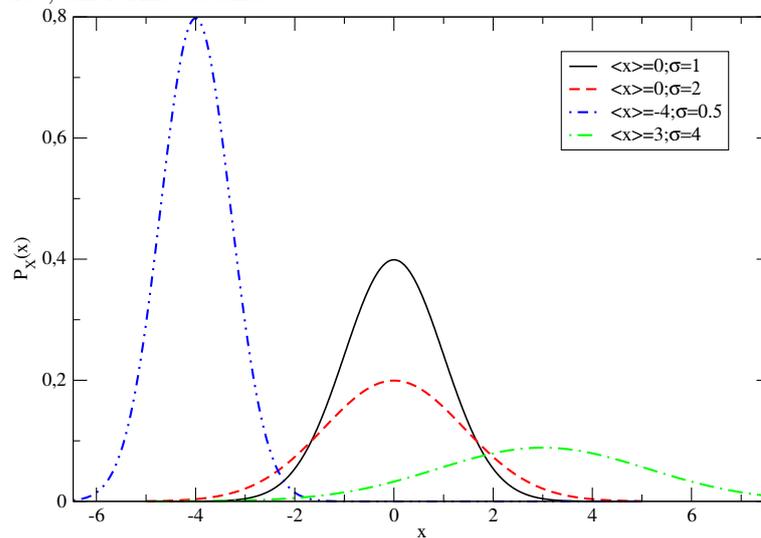
$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - \langle x \rangle)^2}{2\sigma^2} \right\}, \quad (2.69)$$

onde $\langle x \rangle$ é a média de x e σ seu desvio padrão.

A distribuição Gaussiana se torna tão importante e estudada, por sua relação íntima com o teorema do limite central. Nessa seção, na qual a distribuição foi apresentada, já se mostra um exemplo claro de aplicação do teorema, onde uma distribuição, no caso a binomial, converge para a distribuição normal quando o número de realizações é grande. Como em diversos sistemas temos distribuições com variáveis descorrelacionadas e que obedecem a outros critérios estabelecidos no teorema, a distribuição gaussiana está presente em inúmeros âmbitos do conhecimento, desde coisas cotidianas como o lançar de vários dados justos, ao agrupamento de galáxias[11].

A curva da distribuição de probabilidade é bem característica, o formato informalmente é chamado de “sino”, por lembrar a geometria do objeto. Na figura 2.8 são apresentadas algumas curvas seguindo a equação 2.69 com alguns parâmetros que as diferenciam, mas nunca perdendo a característica de ter uma cauda que decresce como uma função exponencial, fato relevante nas análises da curva normal.

Figura 2.8: Alguns exemplos de curvas da distribuição gaussiana com alguns parâmetros diversos. É possível notar que a distribuição é simétrica, possuindo portanto *skewness* nula e tem um formato bastante característico, em forma de sino.



Fonte: Autor, 2014.

Os momentos Gaussianos

1. Média

A média é o primeiro parâmetro que pode ser encarado quando se observa uma distribuição qualquer de probabilidade. No caso da distribuição gaussiana, sua média já está definida dentro da própria fundamentação da equação 2.69. É possível observar isso fazendo uso da equação 2.10:

$$\langle x \rangle = \int_{-\infty}^{\infty} x \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx \quad (2.70)$$

é possível resolver essa equação usando um artifício bem simples, amplamente utilizado em mecânica estatística. Calculando a derivada da função $e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ em relação a μ , o resultado será:

$$\frac{d}{d\mu} \left(e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) = \frac{(x-\mu)}{\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2.71)$$

apenas usando a regra da cadeia. Sendo assim, é fácil notar que:

$$xe^{-\frac{(x-\mu)^2}{2\sigma^2}} = \sigma^2 \frac{d}{d\mu} \left(e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) + \mu e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2.72)$$

logo, a equação 2.70 será então

$$\langle x \rangle = \sigma^2 \frac{d}{d\mu} \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx + \mu \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx \quad (2.73)$$

pela condição de normalização, $\int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx = 1$, portanto,

$$\begin{aligned} \langle x \rangle &= \sigma^2 \frac{d}{d\mu} (1) + \mu \times 1 \\ &= \mu \end{aligned} \quad (2.74)$$

pois a derivada de uma constante é nula. Dessa forma a média em uma distribuição gaussiana é o parâmetro que está subtraindo x dentro da exponencial.

2. O segundo momento e o desvio padrão

Para o segundo momento, utilizando a equação 2.21 é possível calcular o segundo momento.

$$\langle x^2 \rangle = \int_{-\infty}^{\infty} x^2 \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx \quad (2.75)$$

aplicando a mesma metodologia anterior e tomando a derivada em segunda ordem do termo $e^{-\frac{(x-\mu)^2}{2\sigma^2}}$,

$$\begin{aligned} \frac{d^2}{d\mu^2} \left(e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) &= \frac{d}{d\mu} \left(\frac{(x-\mu)}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) \\ &= -\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma^2} + \frac{(x-\mu)^2}{\sigma^4} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\ &= -\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma^2} + \frac{(x^2 - 2x\mu + \mu^2)}{\sigma^4} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \end{aligned} \quad (2.76)$$

sendo assim, é trivial notar que:

$$x^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \sigma^4 \frac{d^2}{d\mu^2} \left(e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) + \sigma^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} + (2x\mu - \mu^2) e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2.77)$$

e portanto a equação 2.75 toma a forma:

$$\begin{aligned} \langle x^2 \rangle &= \sigma^4 \frac{d^2}{d\mu^2} \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx + \sigma^2 \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx + \\ &+ 2\mu \int_{-\infty}^{\infty} x \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx - \mu^2 \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx \\ &= \sigma^4 \frac{d^2}{d\mu^2} (1) + \sigma^2 (1) + 2\mu \langle x \rangle - \mu^2 (1) \\ &= \sigma^2 + \mu^2 \end{aligned} \quad (2.78)$$

sendo esse então o segundo momento estatístico para qualquer curva gaussiana, com parâmetros σ e μ .

É possível também calcular o segundo momento centrado na média, o desvio padrão ao quadrado. As equações 2.22 e 2.25 fornecem a visão geral para o cálculo do momento centrado na média e o específico para o segundo momento, respectivamente. Aplicando então a equação 2.25:

$$\mu'_2 = \langle x^2 \rangle - \langle x \rangle^2 = \sigma^2 + \mu^2 - (\mu)^2 = \sigma^2 \quad (2.79)$$

Revisitando as equações da seção 2.3.2, em especial a equações 2.16 e 2.17, é possível observar que o resultado mostrado na equação 2.79 representa exatamente o desvio padrão da distribuição gaussiana ao quadrado.

Dessa forma, utilizando esses dois resultados, é possível afirmar que a distribuição gaussiana possui dos parâmetros relevantes, μ e σ , que através das definições propostas pelas as equações 2.74 e 2.79 são identificados como a média e o desvio padrão, respectivamente, da distribuição em questão.

3. O terceiro momento e a *skewness*

A metodologia mais uma vez é repetida, aplicando a equação 2.21 é obtido:

$$\langle x^3 \rangle = \int_{-\infty}^{\infty} x^3 \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx \quad (2.80)$$

e portanto, para resolver a integral é calculada a derivada de terceira ordem em μ do termo $e^{-\frac{(x-\mu)^2}{2\sigma^2}}$:

$$\begin{aligned} \frac{d^3}{d\mu^3} \left(e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) &= \frac{d}{d\mu} \left(\frac{d^2}{d\mu^2} \left(e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) \right) \\ &= \frac{d}{d\mu} \left(-\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma^2} + \frac{(x-\mu)^2}{\sigma^4} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) \\ &= \frac{(x-\mu)^3}{\sigma^6} e^{-\frac{(x-\mu)^2}{2\sigma^2}} - \frac{3(x-\mu)}{\sigma^4} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\ &= \frac{(x^3 - 3x^2\mu + 3x\mu^2 - \mu^3)}{\sigma^6} e^{-\frac{(x-\mu)^2}{2\sigma^2}} - \frac{3(x-\mu)}{\sigma^4} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \end{aligned} \quad (2.81)$$

logo, resolvendo a equação para $x^3 e^{-\frac{(x-\mu)^2}{2\sigma^2}}$,

$$x^3 e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \sigma^6 \frac{d^3}{d\mu^3} \left(e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) + 3\sigma^2(x-\mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} + (3x^2\mu - 3x\mu^2 + \mu^3) e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2.82)$$

Dessa forma a equação 2.80 poderá ser escrita como:

$$\begin{aligned} \langle x^3 \rangle &= \sigma^6 \frac{d^3}{d\mu^3} \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx + 3\sigma^2 \int_{-\infty}^{\infty} x \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx - \\ &\quad - 3\mu\sigma^2 \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx + 3\mu \int_{-\infty}^{\infty} x^2 \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx - \\ &\quad - 3\mu^2 \int_{-\infty}^{\infty} x \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx + \mu^3 \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx \end{aligned} \quad (2.83)$$

ou seja, resolvendo as integrais com as definições da equação 2.21

$$\begin{aligned} \langle x^3 \rangle &= \sigma^6 \frac{d^3}{d\mu^3} (1) + 3\sigma^2 \langle x \rangle - 3\mu\sigma^2 (1) + 3\mu \langle x^2 \rangle - 3\mu^2 \langle x \rangle + \mu^3 (1) \\ &= 3\mu\sigma^2 + \mu^3 \end{aligned} \quad (2.84)$$

O terceiro momento da distribuição gaussiana, depende da média e do seu desvio padrão, como era de se esperar, os dois parâmetros relevantes ao sistema. Calculando agora o terceiro momento centrado na média, afim de obter o valor da *skewness*, é possível, através da equação 2.26, verificar que:

$$\begin{aligned} \mu'_3 &= \langle x^3 \rangle - 3\langle x^2 \rangle \langle x \rangle + 2\langle x \rangle^3 \\ &= 3\mu\sigma^2 + \mu^3 - 3(\sigma^2 + \mu^2)\mu + 2\mu^3 \\ &= 0 \end{aligned} \quad (2.85)$$

e por consequência, a *skewness*, que é dada pela equação 2.29, será também nula,

$$S_{gauss} = \frac{\mu'_3}{\sigma^3} = \frac{0}{\sigma^3} = 0 \quad (2.86)$$

corroborando o fato de que a distribuição gaussiana é uma distribuição simétrica com relação a média, pois exatamente o seu grau de assimetria com relação ao parâmetro é nulo.

4. O quarto momento e a *kurtosis*

Para o quarto e último momento a ser abordado nessa tese, se faz necessário o uso da equação 2.21 para obter a integração referente ao quarto momento:

$$\langle x^4 \rangle = \int_{-\infty}^{\infty} x^4 \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx \quad (2.87)$$

Afim de resolver a integral acima, calcula-se a derivada da função $e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ em quarta ordem em μ :

$$\begin{aligned} \frac{d^4}{d\mu^4} \left(e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) &= \frac{d}{d\mu} \left(\frac{d^3}{d\mu^3} \left(e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) \right) \\ &= \frac{d}{d\mu} \left(\frac{(x-\mu)^3}{\sigma^6} e^{-\frac{(x-\mu)^2}{2\sigma^2}} - \frac{3(x-\mu)}{\sigma^4} \right) \\ &= \frac{(x-\mu)^4}{\sigma^8} e^{-\frac{(x-\mu)^2}{2\sigma^2}} + \frac{3}{\sigma^4} e^{-\frac{(x-\mu)^2}{2\sigma^2}} - \frac{6(x-\mu)^2}{\sigma^6} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\ &= \frac{x^4 - 4x^3\mu + 6x^2\mu^2 - 4x\mu^3 + \mu^4}{\sigma^8} e^{-\frac{(x-\mu)^2}{2\sigma^2}} + \\ &+ \frac{3}{\sigma^4} e^{-\frac{(x-\mu)^2}{2\sigma^2}} - \frac{6(x-\mu)^2}{\sigma^6} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \end{aligned} \quad (2.88)$$

resolvendo para $x^4 e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ afim de resolver a equação 2.87,

$$\begin{aligned} x^4 e^{-\frac{(x-\mu)^2}{2\sigma^2}} &= \sigma^8 \frac{d^4}{d\mu^4} \left(e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) + (4x^3\mu - 6x^2\mu^2 + 4x\mu^3 - \mu^4) e^{-\frac{(x-\mu)^2}{2\sigma^2}} - \\ &- 3\sigma^4 e^{-\frac{(x-\mu)^2}{2\sigma^2}} + 6\sigma^2(x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} \end{aligned} \quad (2.89)$$

por consequência a equação 2.87 será dada por:

$$\begin{aligned} \langle x^4 \rangle &= \sigma^8 \frac{d^4}{d\mu^4} \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx + 4\mu \int_{-\infty}^{\infty} x^3 \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx - 6\mu^2 \int_{-\infty}^{\infty} x^2 \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx + \\ &+ 4\mu^3 \int_{-\infty}^{\infty} x \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx - \mu^4 \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx - 3\sigma^4 \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx + \\ &+ 6\sigma^2 \int_{-\infty}^{\infty} x^2 \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx - 12\sigma^2\mu \int_{-\infty}^{\infty} x \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx + 6\sigma^2\mu^2 \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx \end{aligned} \quad (2.90)$$

logo,

$$\begin{aligned}\langle x^4 \rangle &= \sigma^8 \frac{d^4}{d\mu^4}(1) + 4\mu \langle x^3 \rangle - 6\mu^2 \langle x^2 \rangle + 4\mu^3 \langle x \rangle - \mu^4(1) - 3\sigma^4(1) + \\ &+ 6\sigma^2 \langle x^2 \rangle - 12\sigma^2 \mu \langle x \rangle + 6\sigma^2 \mu^2(1) \\ &= 3\sigma^4 + 6\sigma^2 \mu^2 + \mu^4\end{aligned}\quad (2.91)$$

Com essa condição, o quarto momento centrado na média, fundamental para o cálculo da *kurtosis*, dado pela equação 2.27, será:

$$\begin{aligned}\mu'_4 &= \langle x^4 \rangle - 4\langle x^3 \rangle \langle x \rangle + 6\langle x^2 \rangle \langle x \rangle^2 - 3\langle x \rangle^4 \\ &= 3\sigma^4 + 6\sigma^2 \mu^2 + \mu^4 - 4(3\mu\sigma^2 + \mu^3)\mu + 6(\sigma^2 + \mu^2)\mu^2 - 3\mu^4 \\ &= 3\sigma^4\end{aligned}\quad (2.92)$$

por sua vez, a *kurtosis*, por sua própria definição (equação 2.30) será para a gaussiana:

$$\kappa = \frac{\mu'_4}{\sigma^4} = \frac{3\sigma^4}{\sigma^4} = 3\quad (2.93)$$

ou seja, tomando a *kurtosis* excessiva, subtraindo 3, a gaussiana passa a ter uma *kurtosis*:

$$\kappa'_{gaussiana} = 3 - 3 = 0\quad (2.94)$$

Esses dois resultados, para a *skewness* e *kurtosis* da curva gaussiana são muito importantes pelo fato de que eles são representados por uma quantidade que não depende dos parâmetros μ ou σ , e por isso, representam um dos testes confiáveis de convergência porque são comuns a qualquer curva gaussiana. É fato que esses testes comparativos de valores gaussianos para *skewness* e *kurtosis*, por si só não garantem que uma distribuição qualquer seja gaussiana, mas são parâmetros que se tornam relevantes por acrescentarem mais informações estatísticas de comparação preciosas à essas análises de convergência ao comportamento gaussiano.

2.5.3 Distribuições α -estáveis de Lévy

Em 1925, o matemático francês Paul Lévy escreveu uma distribuição de probabilidades mais geral[15]. As distribuições de Lévy formam um conjunto dependente do parâmetro α que ajustado fornece diversas distribuições.

As distribuições α -estáveis de Lévy são definidas a partir da função característica que é dada por[19]:

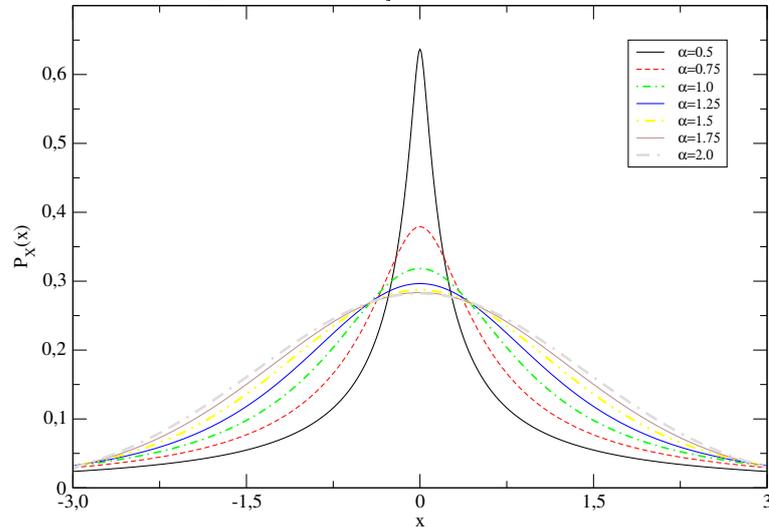
$$f_x(k) = \begin{cases} \exp \{ ik\mu - |ck|^\alpha [1 - i\beta \operatorname{sgn}(k) \operatorname{tg}(\pi\alpha/2)] \} & \text{se } \alpha \neq 1 \\ \exp \{ ik\mu - |ck|^\alpha [1 + i\beta \operatorname{sgn}(k)(2 \log(k)/\pi)] \} & \text{se } \alpha = 1 \end{cases}\quad (2.95)$$

onde α está no intervalo $(0, 2]$ e é o parâmetro característico das distribuições, sendo relevante em sua configuração geométrica e em sua cauda assintótica, β está entre $[-1, 1]$ e é relacionado com a assimetria da distribuição, μ em algumas distribuições é relacionado com a média, o parâmetro $c > 0$ é a sua escala e a função $\operatorname{sgn}(k)$ é a função sinal e é definida pela seguinte proposição:

$$\operatorname{sgn}(k) := \begin{cases} -1 & \text{se } k < 0, \\ 0 & \text{se } k = 0, \\ 1 & \text{se } k > 0. \end{cases}\quad (2.96)$$

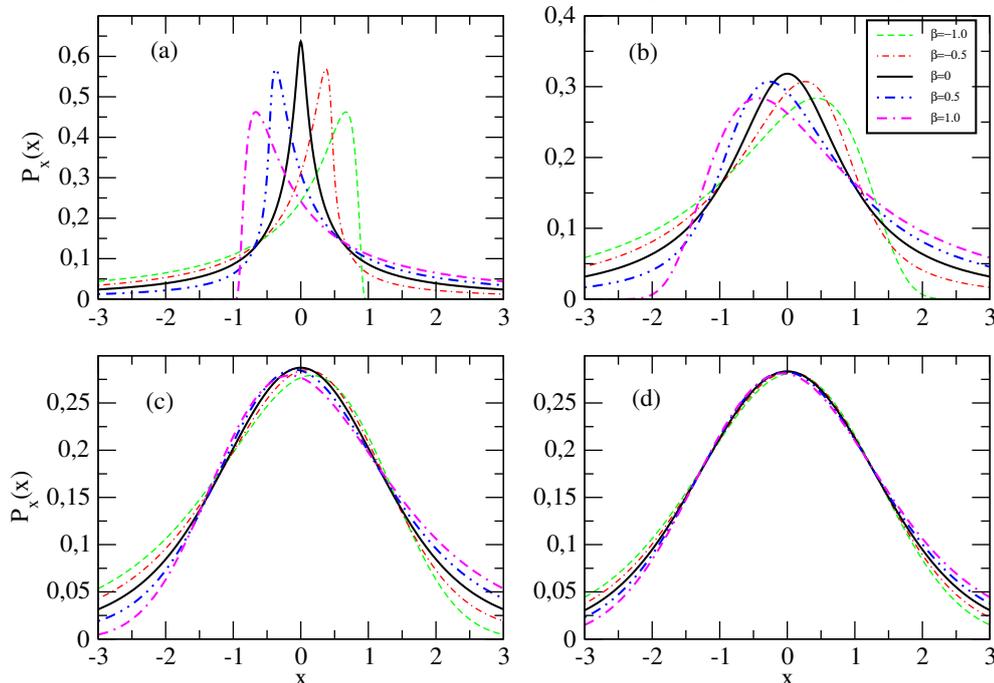
É possível observar nas figuras 2.9 e 2.10 apenas um apanhado mínimo de possibilidades que as distribuições α -estáveis de Lévy possuem, mostrando a relação que α e β tem com curva geométrica apresentada por cada distribuição particular.

Figura 2.9: Algumas distribuições provenientes da equação 2.95 para diferentes α e com $\beta = \mu = 0$ e com $c = 1$. Dados gerados usando o programa R[20]



Fonte: Autor, 2014.

Figura 2.10: Algumas distribuições provenientes da equação 2.95 com $\mu = 0$, $c = 1$, (a) $\alpha = 0,5$, (b) $\alpha = 1,0$, (c) $\alpha = 1,5$, (d) $\alpha = 1,75$ e variados valores de β . Dados gerados usando o programa R[20]

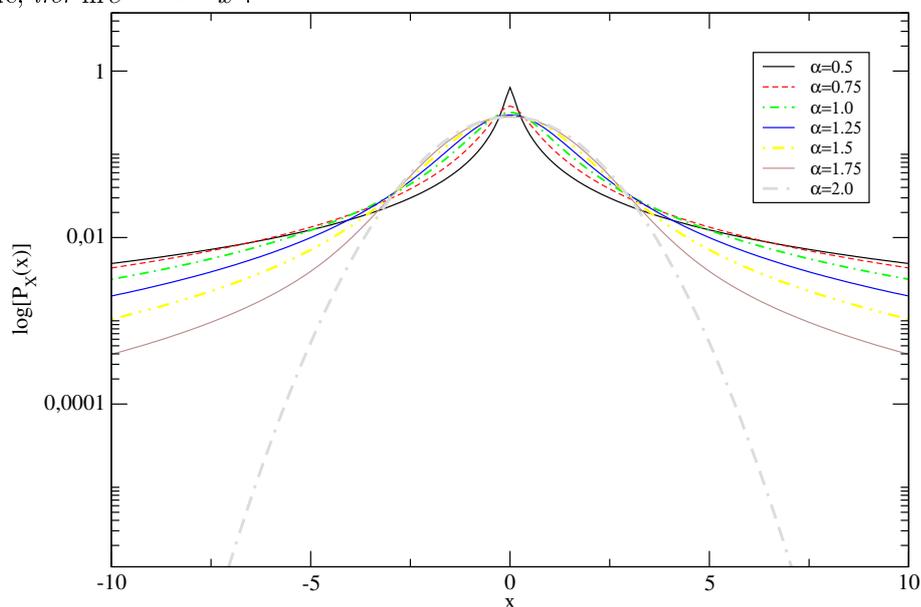


Fonte: Autor, 2014.

É possível observar que as distribuições para $\alpha < 2$ a sua cauda, ou seja, o seu comportamento assintótico, não aparenta uma curva exponencial e sim uma cauda como lei de potência. Essa probabilidade não nula para eventos extremos, pode ser melhor

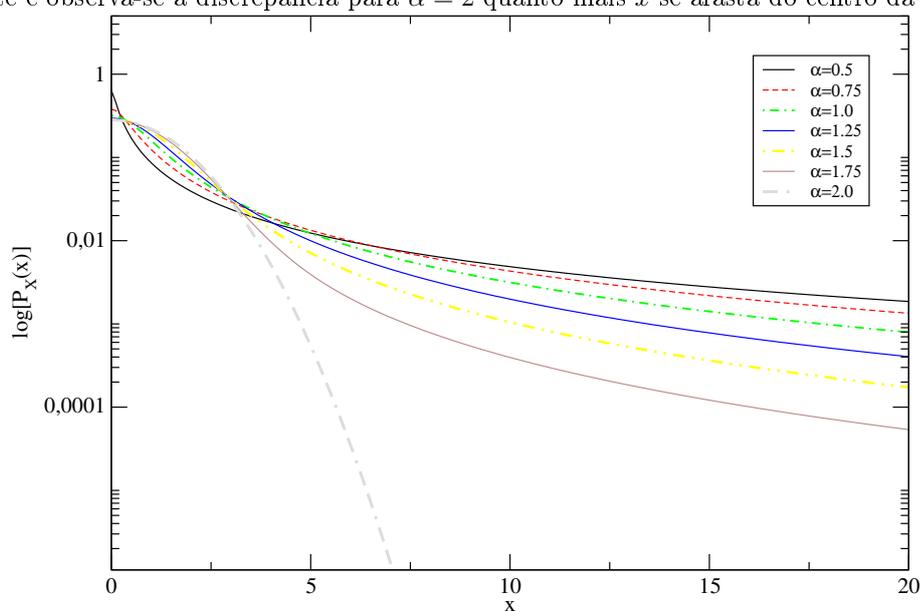
observada em um eixo semi-log, como mostram as figuras 2.11 e 2.12. O comportamento exponencial pode ser observado na cauda para $\alpha = 2$ onde a queda é muito mais acentuada.

Figura 2.11: Eixo em semi-log da figura 2.9, mostrando como o decaimento sugere uma lei de potência, indicado pelo comportamento não polinomial das distribuições quando sujeitas ao logaritmo da probabilidade, *i.e.* $\ln e^{-x^2} = -x^2$.



Fonte: Autor, 2014.

Figura 2.12: Figura voltada à cauda, onde o comportamento em lei de potência para $\alpha < 2$ se torna mais evidente e observa-se a discrepância para $\alpha = 2$ quanto mais x se afasta do centro da distribuição.



Fonte: Autor, 2014.

Para $\alpha < 2$ no regime assintótico, quando x é muito maior, ou menor, que a média

ou seja, na cauda da distribuição, o seu comportamento é tipo lei de potência, e dado como[19]:

$$P_X(x) \sim \frac{\alpha^\alpha(1+\beta)\text{sen}(\pi\alpha/2)\Gamma(\alpha)/\pi}{|x|^{1+\alpha}}. \quad (2.97)$$

ou seja, a cauda é proporcional a

$$P_X(x) \sim \frac{1}{|x|^{1+\alpha}}. \quad (2.98)$$

Esse comportamento em forma de lei de potência na cauda, permite que essas distribuições possam mapear problemas que são formados por eventos extremos, ou seja, ocorrências onde a variável aleatória x toma valores muito fora da média.

As distribuições α -estáveis de Lévy tentam generalizar um conjunto de distribuições estáveis, àquelas que possuem a propriedade de que a combinação linear de duas variáveis da mesma distribuição resultam nessa mesma distribuição, alterando apenas alguns parâmetros.

Apesar de uma classe bem geral de distribuições, a equação 2.95 possui apenas 3 soluções analíticas. Para $\alpha = 0.5$ e $\beta = 1$, a solução é chamada de distribuição de Lévy-Smirnov. Com $\alpha = 1$ e $\beta = 0$, a solução é a distribuição de Cauchy, ou por vezes chamada de Lorentziana. Por fim, e não menos importante, para $\alpha = 2$, a distribuição resultante da transformada de Fourier é a distribuição Gaussiana. É possível demonstrar cada uma dessas soluções analíticas e suas propriedades com o conhecimento de cálculo.

A função característica para a distribuição de Lévy-Smirnov parte da função característica das distribuições α -estáveis de Lévy, com $\alpha = 0.5$ e $\beta = 1$, como afirmado. Dessa forma a função no espaço de momentos será:

$$f_x(k) = e^{ik\mu - \sqrt{-2ick}} \quad (2.99)$$

Sendo assim, por definição, a função densidade de probabilidade é dada pela transformada inversa de Fourier, e portanto toma a forma:

$$\begin{aligned} P(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik\mu - \sqrt{-2ick}} e^{-ikx} dk & (2.100) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik\mu - \sqrt{-2ick} - ikx} dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ik(x-\mu) - \sqrt{-2ick}} dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-[ik(x-\mu) + \sqrt{-2ick}]} dk \end{aligned}$$

definindo:

$$y = i(x - \mu) \quad (2.101)$$

$$z = -2ic \quad (2.102)$$

a integral se torna visualmente mais simplificada:

$$P(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-(ky + \sqrt{kz})} dk \quad (2.103)$$

afim de simplificar a integral mais uma vez, é possível fazer o seguinte rearranjo matemático:

$$ky + \sqrt{kz} = (\sqrt{ky})^2 + 2(\sqrt{ky})\left(\sqrt{\frac{z}{4y}}\right) + \left(\sqrt{\frac{z}{4y}}\right)^2 - \left(\sqrt{\frac{z}{4y}}\right)^2 \quad (2.104)$$

$$= \left(\sqrt{ky} + \sqrt{\frac{z}{4y}}\right)^2 - \frac{z}{4y} \quad (2.105)$$

Com esse rearranjo a equação 2.103 será dada por:

$$P(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp \left\{ - \left[\left(\sqrt{ky} + \sqrt{\frac{z}{4y}} \right)^2 - \frac{z}{4y} \right] \right\} dk \quad (2.106)$$

$$P(x) = \frac{e^{\frac{z}{4y}}}{2\pi} \int_{-\infty}^{\infty} e^{-(\sqrt{ky} + \sqrt{\frac{z}{4y}})^2} dk \quad (2.107)$$

Para resolver essa integral, é possível fazer uma mudança de variável que considere a solução positiva, ou seja:

$$-u = \sqrt{ky} \Rightarrow k = \frac{u^2}{y} \rightarrow dk = \frac{2u}{y} du \quad (2.108)$$

$$\theta = \sqrt{\frac{z}{4y}} \quad (2.109)$$

e com isso a equação 2.107 será dada por:

$$P(x) = \frac{e^{\theta^2}}{y\pi} \int_{-\infty}^{\infty} u e^{-(u+\theta)^2} du \quad (2.110)$$

que pode ser reorganizada de uma forma mais familiar:

$$P(x) = \frac{e^{\theta^2}}{y\pi} \int_{-\infty}^{\infty} \frac{\sqrt{2\pi(\sqrt{1/2})^2} u e^{-\frac{(u+\theta)^2}{2(\sqrt{1/2})^2}}}{\sqrt{2\pi(\sqrt{1/2})^2}} du \quad (2.111)$$

$$= \frac{e^{\theta^2} \sqrt{\pi}}{y\pi} \int_{-\infty}^{\infty} \frac{u e^{-\frac{(u+\theta)^2}{2(\sqrt{1/2})^2}}}{\sqrt{2\pi(\sqrt{1/2})^2}} du \quad (2.112)$$

$$= \frac{e^{\theta^2}}{y\sqrt{\pi}} \langle u \rangle \quad (2.113)$$

observando as equações gaussianas, é possível notar que $\langle u \rangle = \theta$, logo

$$P(x) = \frac{e^{\theta^2}}{y\sqrt{\pi}} (\theta) \quad (2.114)$$

$$= \frac{\theta e^{\theta^2}}{y\sqrt{\pi}} \quad (2.115)$$

Pela equação 2.109, a função será:

$$P(x) = \frac{e^{\frac{z}{4y}}}{y\sqrt{\pi}} \sqrt{\frac{z}{4y}} \quad (2.116)$$

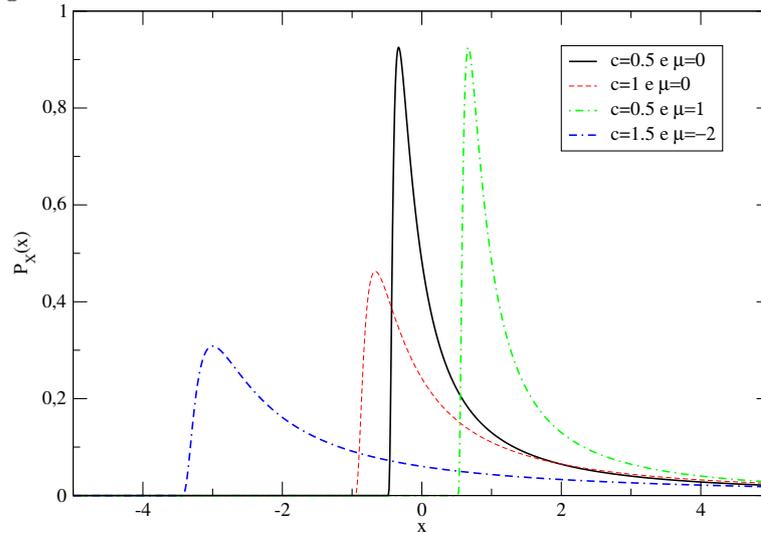
e usando as equações 2.101 e 2.102, é possível retornar as variáveis originais:

$$\begin{aligned}
 P(x) &= \frac{e^{\frac{-2ic}{4i(x-\mu)}}}{[i(x-\mu)]\sqrt{\pi}} \sqrt{\frac{-2ic}{4i(x-\mu)}} \\
 &= \frac{e^{\frac{-c}{2(x-\mu)}}}{[i(x-\mu)]\sqrt{\pi}} \sqrt{\frac{-c}{2(x-\mu)}} \\
 &= \frac{e^{\frac{-c}{2(x-\mu)}}}{[i(x-\mu)]\sqrt{\pi}} i \sqrt{\frac{c}{2(x-\mu)}} \\
 &= \frac{e^{\frac{-c}{2(x-\mu)}}}{(x-\mu)\sqrt{\pi}} \sqrt{\frac{c}{2(x-\mu)}} \\
 &= \sqrt{\frac{c}{2\pi}} \frac{e^{\frac{-c}{2(x-\mu)}}}{(x-\mu)^{3/2}}
 \end{aligned} \tag{2.117}$$

Formalizando a função densidade de probabilidade da distribuição de Lévy-Smirnov.

Na figura 2.13 pode-se verificar algumas curvas que seguem a função dada pela equação 2.117, mostrando a característica de cauda como lei de potência e também a sua clara assimetria, provocada pela alteração do fator β , gerando diversas distribuições.

Figura 2.13: A distribuição de Lévy-Smirnov para alguns parâmetros



Fonte: Autor, 2014.

Avançando os valores de α onde solução da função característica é analítica, surge a distribuição de Cauchy, ou algumas vezes conhecida como distribuição de Lorentz, ou Lorentziana. Para essa distribuição $\alpha = 1$ e $\beta = 0$. A função característica então será dada por:

$$f_x(k) = e^{ik\mu - |ck|} = \begin{cases} e^{ik\mu - ck} & \text{se } k > 0 \\ e^{ik\mu + ck} & \text{se } k < 0 \end{cases} \tag{2.118}$$

Tomando então a transformada de Fourier afim de determinar a função densidade de

probabilidade:

$$\begin{aligned}
P(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik\mu - |ck|} e^{-ikx} dk \\
&= \frac{1}{2\pi} \int_0^{\infty} e^{ik\mu - ck} e^{-ikx} dk + \frac{1}{2\pi} \int_{-\infty}^0 e^{ik\mu + ck} e^{-ikx} dk \\
&= \frac{1}{2\pi} \int_0^{\infty} e^{-k[i(x-\mu)+c]} dk + \frac{1}{2\pi} \int_{-\infty}^0 e^{k[-i(\mu-x)+c]} dk \\
&= \frac{1}{2\pi} \left[\frac{-e^{-k[i(x-\mu)+c]}}{[i(x-\mu)+c]} \right]_0^{\infty} + \frac{1}{2\pi} \left[\frac{e^{k[-i(\mu-x)+c]}}{[-i(\mu-x)+c]} \right]_{-\infty}^0 \\
&= \frac{1}{2\pi[i(x-\mu)+c]} + \frac{1}{2\pi[-i(\mu-x)+c]} \\
&= \frac{i(x-\mu)+c - i(x-\mu)+c}{2\pi[i(x-\mu)+c][-i(x-\mu)+c]} \\
&= \frac{2c}{2\pi\{-[i(x-\mu)]^2 + c^2\}}
\end{aligned} \tag{2.119}$$

Portanto, a distribuição de Cauchy tem como função densidade de probabilidade a seguinte equação:

$$P(x) = \frac{c}{\pi[(x-\mu)^2 + c^2]} \tag{2.120}$$

Na figura 2.14 é possível observar algumas das várias faces que a distribuição de Cauchy apresenta. É interessante observar que em todas elas a cauda da distribuição é uma cauda grossa, ou seja, com uma queda que não se escreve como exponencial e sim como lei de potência. Esse comportamento fica melhor observável ao tomarmos o logaritmo natural da probabilidade, pois se a distribuição for dada por uma exponencial, seja do tipo $P_X(x) = Ae^{x^n}$, o logaritmo será: $\ln P_X(x) = \ln A + x^n$, e portanto um gráfico de $\ln P_X(x)$ por x , será um polinômio, enquanto se a cauda for dada por uma lei de potência, o gráfico remete ao logaritmo de x .

Por fim, entre as soluções analíticas da função característica, para $\alpha = 2$ surge a curva gaussiana. Partindo da função característica com esse parâmetro:

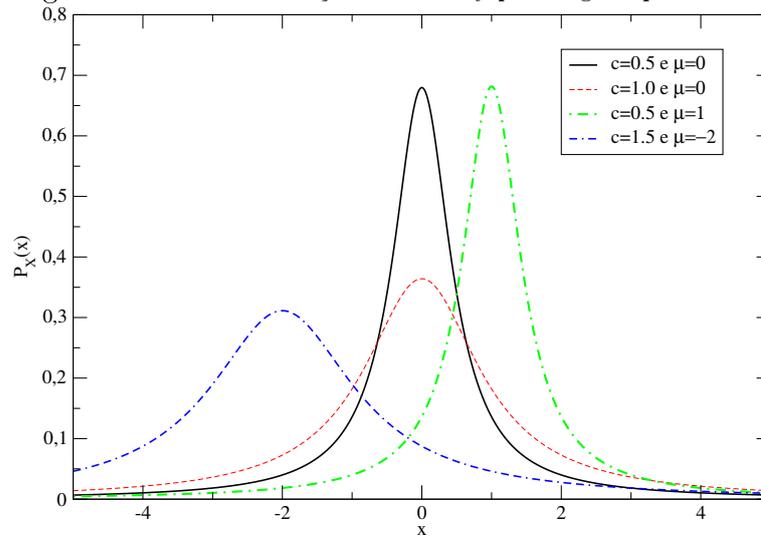
$$f_x(k) = e^{ik\mu - (ck)^2} = e^{ik\mu - c^2k^2} \tag{2.121}$$

Dessa forma a função densidade de probabilidade gaussiana, que é dada pela transformada de Fourier da equação será portanto:

$$\begin{aligned}
P(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik\mu - c^2k^2} e^{-ikx} dk \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ik(x-\mu) - c^2k^2} dk \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-[ik(x-\mu) + c^2k^2]} dk
\end{aligned} \tag{2.122}$$

Logo, é possível completar o quadrado dentro da exponencial afim de resolver a integral, como foi feito anteriormente com a distribuição de Lévy-Smirnov. Sendo assim, com

Figura 2.14: A distribuição de Cauchy para alguns parâmetros



Fonte: Autor, 2014.

$$y = i(x - \mu):$$

$$\begin{aligned} ky + c^2k^2 &= ky + c^2k^2 + \left(\frac{y}{2c}\right)^2 - \left(\frac{y}{2c}\right)^2 \\ &= \left(ck + \frac{y}{2c}\right)^2 - \left(\frac{y}{2c}\right)^2 \end{aligned} \quad (2.123)$$

Então, a equação 2.122 com a reformulação proposta na equação 2.123, será:

$$\begin{aligned} P(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\left[\left(ck + \frac{y}{2c}\right)^2 - \left(\frac{y}{2c}\right)^2\right]} dk \\ &= \frac{e^{\left(\frac{y}{2c}\right)^2}}{2\pi} \int_{-\infty}^{\infty} e^{-\left(ck + \frac{y}{2c}\right)^2} dk \\ &= \frac{e^{\left(\frac{y}{2c}\right)^2}}{2\pi} \frac{\sqrt{\pi}}{c} \end{aligned} \quad (2.124)$$

Como $y = i(x - \mu)$, logo:

$$\begin{aligned} P(x) &= \frac{e^{\left(\frac{i(x-\mu)}{2c}\right)^2}}{2\pi} \frac{\sqrt{\pi}}{c} \\ &= \frac{e^{-\frac{(x-\mu)^2}{4c^2}}}{2\pi} \frac{\sqrt{\pi}}{c} \end{aligned} \quad (2.125)$$

reorganizando os fatores, afim de mostrar uma equação similar à 2.69:

$$P(x) = \frac{e^{-\frac{(x-\mu)^2}{2(\sqrt{2}c)^2}}}{\sqrt{2\pi(\sqrt{2}c)^2}} \quad (2.126)$$

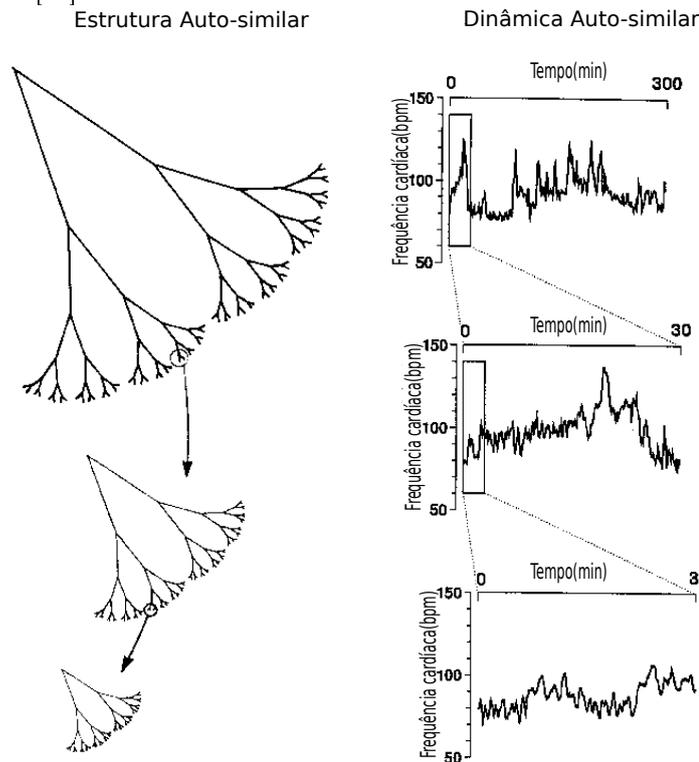
onde é trivial notar que $\sigma = \sqrt{2}c$. Algumas das diferentes formas que a distribuição gaussiana pode ter, são observadas na figura 2.8, apresentada anteriormente nessa mesma tese.

2.6 Fractalidade em Séries temporais

Uma propriedade muito interessante, presente em diversos sistemas matemáticos, biológicos, econômicos e algumas séries temporais[21, 22, 23, 24], a auto-similaridade. Tais sistemas, quando tem sua escala alterada apresentam o mesmo padrão, ou seja, esses sistemas possuem uma invariância por escala, podendo ser discreta ou contínua. Essa é uma propriedade típica de sistemas chamados fractais.

Além da auto-similaridade, fractais também apresentam dimensão topológica presente no conjunto dos reais, diferente de outros padrões geométricos que possuem sua dimensão em um dado número natural.

Figura 2.15: Representação ilustrativa da auto-similaridade presente em figuras geométricas e séries temporais. Adaptada de [25]



Fonte: Golberger, 1996 [25]

A auto-similaridade apresenta algumas limitações em casos reais, pois estruturas reais, não são infinitas, e portanto não podem se comportar como infinitas. Além disso é possível que o sistema tome uma característica muito distinta quando a escala é arbitrariamente alterada, de tal forma que os padrões antes existentes, não possam mais ser observados. Um exemplo são os galhos de algumas árvores e suas ramificações, a auto-similaridade é apresentada nesse contexto, mas ao atingir o nível molecular, esse padrão usualmente se perde, tomando um outro formado específico[26, 27].

Para uma série temporal, o principal objeto de estudo dessa tese, define-se como uma série auto-similar caso a mesma satisfizer a propriedade de invariância por escala. Esse contexto pode ser escrito como:

$$X(t) \stackrel{d}{\equiv} a^\kappa X\left(\frac{t}{a}\right) \quad (2.127)$$

onde, $X(t)$ é uma série de variáveis aleatórias x , que são organizadas em uma sequência cronológica, $\stackrel{d}{\equiv}$ implica que as propriedades estatísticas de ambos os lados da igualdade são estatisticamente similares ou idênticas, ou seja, as distribuições tem aproximadamente ou exatamente a mesma média, variância, *skewness*, *kurtosis* e outros momentos estatísticos, a é um parâmetro qualquer utilizado arbitrariamente para alterar a escala da série e o expoente κ é denominado parâmetro de auto-similaridade da série temporal.

A definição formalizada implica que uma série temporal será auto-similar se e somente se, os dados da série, as variáveis aleatórias x for reescalado na forma: $x \rightarrow a^\kappa x$, e o eixo temporal, que define a amostragem, for reescalado por um fator a ($t \rightarrow \frac{t}{a}$), e a série resultante possuir as mesmas propriedades estatísticas que a original.

Em 1951, Hürst, um hidrólogo inglês, estudou a série temporal da capacidade armazenamento de água vinda do Rio Nilo no Egito e observou que existia uma clara dependência temporal segundo uma lei de potência. Hürst, com dados empíricos, que dado uma quantidade $Q(t)$ de volume de água, um recipiente deverá ter uma amplitude $R(t)$. Se o desvio padrão da série temporal for $S(t)$, Hürst observou que a razão R/S era proporcional a t^H , onde H , ele mediu como sendo 0,7[28]. Esse conceito usado por Hurst para o cálculo do expoente, ficou conhecido como método R/S .

Leis de potência são funções que estão diretamente relacionadas com algumas propriedades fractais, é fácil notar que uma função $f(x) = x^n$ possui características auto-similares, pois se x for reescalado por um fator a , é trivial que: $f(ax) = (ax)^n = a^n x^n = a^n f(x)$.

Nesse contexto o expoente de Hürst, H , foi utilizado em análises de outras séries temporais[29, 30] e um outro método foi desenvolvido, afim de eliminar algumas tendências que provocavam erros no método R/S , o método DFA[31]. O método DFA (*Detrend Fluctuation Analysis*) consiste em fornecer o expoente de Hürst relacionado a uma dada série temporal.

Dada uma série temporal, o primeiro passo para a análise DFA é integrar a série temporal, $x(t)$ com N pontos, ou seja, o tempo total será N , obtendo a função:

$$y(t) = \sum_{t'}^t (x_{t'} - \langle x \rangle); \quad (2.128)$$

onde temos $t = 1, \dots, N$.

Com a nova série, $y(t)$, basta dividi-la em N_τ intervalos temporais não superpostos de tamanho τ . Através de uma regressão do n -ésimo segmento ($n = 1, \dots, N_\tau$), é determinado o melhor ajuste para a tendência local da série $y_n(i)$, $i = 1, \dots, \tau$. Por definição o segundo momento para o n -ésimo segmento é dado por:

$$F_2(n, \tau) \equiv \frac{1}{\tau} \sum_{i=1}^{\tau} |y((n-1)\tau + i) - y_n(i)|^2. \quad (2.129)$$

Portanto a flutuação para a série completa se escreve:

$$F_2(\tau) = \frac{1}{N_\tau} \sum_{n=1}^{N_\tau} F_2(n, \tau). \quad (2.130)$$

E dessa forma, fica claro o aparecimento de uma lei de potência envolvendo o parâmetro τ , que possui dimensão de tempo, e o expoente de Hürst, na forma:

$$F_2(\tau) \sim \tau^{2H} \quad (2.131)$$

Essa é uma análise mais sensível das flutuações nas variáveis da série e sem tendências, restrita apenas pela quantidade de dados, que pode indicar uma estatística pobre e resultados aquém da realidade.

O expoente de Hürst também é relacionado com a dimensão fractal do sistema, d , de tal forma que:

$$d = 2 - H \quad (2.132)$$

Para alguns sistemas, no entanto, se faz necessária uma análise de um espectro contínuo de expoentes de escala, generalizando assim o expoente de Hürst. Tais sistemas usualmente estão próximos à transições de fase, ou são relacionados à correlações de longo alcance inerentes ao sistema[32]. Esses sistemas são chamados de multifractais e para essa análise é definido o expoente de Hürst generalizado, $H(q)$, depende do q -ésimo momento, onde a notação de q como momento, é corriqueira na literatura desse sistema. O método DFA agora é generalizado ao método MF-DFA (*Multifractal Detrended Fluctuation Analysis*), de tal forma que a flutuação é definida para o q -ésimo momento como sendo:

$$F_q(\tau) = \frac{1}{N_\tau} \sum_{n=1}^{N_\tau} [F_2(n, \tau)]^{q/2}, \quad (2.133)$$

e a flutuação escala com:

$$F_q(\tau) \sim \tau^{qh(q)}, \quad (2.134)$$

Essa análise explora os valores de q pertencentes ao conjunto dos números reais e faz com que grandes flutuações se sobressaiam para q positivo, e as pequenas flutuações sejam contribuídas para momentos negativos.

Outra técnica utilizada para determinar as propriedades multifractais do sistema é o cálculo do espectro de singularidades, $f(\alpha)$, onde α é o índice de singularidade, ou também chamado de expoente de Hölder [33, 34, 35]. A dimensão multifractal generalizada é dada por:

$$D_q = \frac{1}{q-1} [q\alpha(q) - f(\alpha(q))]. \quad (2.135)$$

de tal forma que é possível calcular que:

$$\frac{d}{dq} [(1-q)D_q] = \alpha, \quad (2.136)$$

usando o postulado que $f'(\alpha) = q$. Seja portanto

$$\beta(q) = (1-q)D_q \quad (2.137)$$

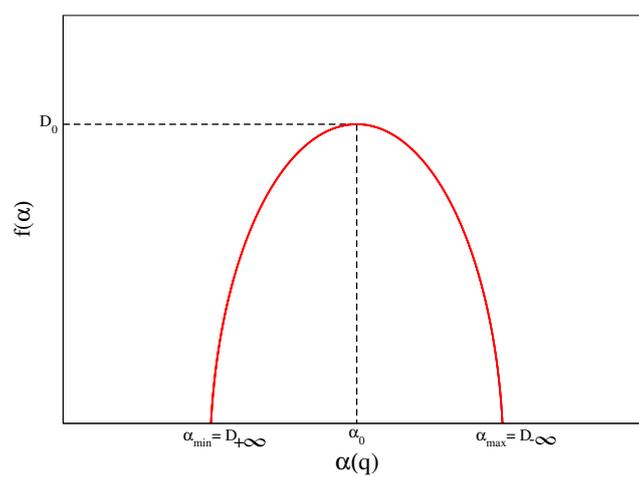
logo é possível escrever que:

$$f(\alpha) = q\alpha - \beta(q). \quad (2.138)$$

é possível relacionar β com o expoente de Hürst segundo[36]:

$$\beta(q) = qh(q) - 1. \quad (2.139)$$

Dessa maneira, é possível determinar via MF-DFA o expoente de Hürst e em seguida a função do espectro de singularidade, $f(\alpha)$. A figura 2.16 mostra o padrão presente nas análises multifractais do espectro multifractal. Quando mais largo se mostra a curva, maior o intervalo de expoentes críticos presentes no sistema, caso a curva tenda a uma largura nula, o sistema tem apenas um expoente de Hölder, e portanto um expoente de Hürst apenas e por isso é um sistema monofractal, ao invés de multifractal.

Figura 2.16: Exemplo do comportamento de $f(\alpha)$ em relação a α . Adaptada de[34]

Fonte: Ott, 1993 [34]

3

O Teorema do Limite Central e as Seções de Lévy

No presente capítulo será apresentada as bases pertinentes ao foco principal da tese, o Teorema do Limite Central em sua forma clássica e uma versão generalizada do mesmo, o Teorema das Seções de Lévy. Os teoremas serão anunciados e seguidos por suas respectivas provas formalizadas.

3.1 Teorema do Limite Central(TLC)

O teorema do limite central em sua forma clássica foi amplamente discutido por Laplace no século XVII e por outros diversos pensadores da área de estatística e matemática. O teorema do limite central afirma que:

Teorema 1 *Seja um conjunto $C_n = X_1; X_2; \dots; X_n$ com n variáveis aleatórias independentes e igualmente distribuídas(i.i.d.), com média e variâncias finitas, o desvio da média aritmética tem uma distribuição que se aproxima da curva normal, ou Gaussiana, a medida que n tende ao infinito.*

Pode-se fazer uma releitura do teorema definindo S_n como o desvio da média na forma:

$$S_n = \frac{X_1 + X_2 + \dots + X_n}{n} - \langle X \rangle = \sum_i^n \frac{(X_i - n\langle X \rangle)}{n} = \sum_i^n s_i, \quad (3.1)$$

onde $\frac{X_1+X_2+\dots+X_n}{n} \neq \langle X \rangle$, pois $\langle X \rangle$ é o valor médio em infinitas realizações e n é um número finito de realizações. O teorema 1 diz que:

$$\lim_{n \rightarrow \infty} P_X(S_n) = \sqrt{\frac{n}{2\pi\sigma_X^2}} e^{-\frac{nS_n^2}{2\sigma_X^2}} \quad (3.2)$$

Com o teorema apresentado, segue uma das provas elegantes do mesmo, via transformada inversa de Fourier, que muito se parece com a prova da lei fraca dos grandes números[18], outro importante resultado em estatística, mas que não será tratado aqui imediatamente.

Prova do Teorema Considere a transformada inversa de Fourier

$$\begin{aligned}
 F^{-1}[P_X(S_n)] &= \int_{-\infty}^{\infty} e^{2i\pi S_n X} P(X) dX \\
 &= \int_{-\infty}^{\infty} \sum_{l=0}^{\infty} \frac{(2i\pi S_n X)^l}{l!} P(X) dX \\
 &= \sum_{l=0}^{\infty} \frac{(2i\pi S_n)^l}{l!} \int_{-\infty}^{\infty} X^l P(X) dX \\
 &= \sum_{l=0}^{\infty} \frac{(2i\pi S_n)^l}{l!} \langle X^l \rangle
 \end{aligned} \tag{3.3}$$

definindo que:

$$\begin{aligned}
 \langle X^l \rangle &= \left\langle \left(\frac{X_1 + X_2 + \dots + X_n}{n} \right)^l \right\rangle \\
 &= \langle n^{-l} (X_1 + X_2 + \dots + X_n)^l \rangle \\
 &= \int_{-\infty}^{\infty} n^{-l} (X_1 + X_2 + \dots + X_n)^l P(X_1) \dots P(X_n) dX_1 \dots dX_n
 \end{aligned} \tag{3.4}$$

Logo da equação 3.3 temos:

$$\begin{aligned}
 F^{-1}[P_X(S_n)] &= \sum_{l=0}^{\infty} \frac{(2i\pi S_n)^l}{l!} \langle X^l \rangle \\
 &= \sum_{l=0}^{\infty} \frac{(2i\pi S_n)^l}{l!} \int_{-\infty}^{\infty} n^{-l} (X_1 + X_2 + \dots + X_n)^l P(X_1) \dots P(X_n) dX_1 \dots dX_n \\
 &= \int_{-\infty}^{\infty} \sum_{l=0}^{\infty} \frac{\left[\frac{(2i\pi S_n (X_1 + X_2 + \dots + X_n))}{n} \right]^l}{l!} P(X_1) \dots P(X_n) dX_1 \dots dX_n \\
 &= \int_{-\infty}^{\infty} e^{\left[\frac{2i\pi S_n (X_1 + X_2 + \dots + X_n)}{n} \right]} P(X_1) \dots P(X_n) dX_1 \dots dX_n \\
 &= \left[\int_{-\infty}^{\infty} e^{2i\pi S_n X_1/n} P(X_1) dX_1 \right] \dots \left[\int_{-\infty}^{\infty} e^{2i\pi S_n X_n/n} P(X_n) dX_n \right]
 \end{aligned}$$

$$\begin{aligned}
 F^{-1}[P_X(S_n)] &= \left[\int_{-\infty}^{\infty} e^{2i\pi S_n X/n} P(X) dX \right]^n \\
 &= \left[\int_{-\infty}^{\infty} \sum_{l=0}^{\infty} \frac{(2i\pi S_n X/n)^l}{l!} P(X) dX \right]^n \\
 &= \left\{ \int_{-\infty}^{\infty} \left[1 + \left(\frac{2i\pi S_n}{n} \right) X + \frac{1}{2} \left(\frac{2i\pi S_n}{n} \right)^2 X^2 + \dots \right] P(X) dX \right\}^n \\
 &= \left[1 + \frac{2i\pi S_n}{n} \langle X \rangle - \frac{1}{2} \left(\frac{2\pi S_n}{n} \right)^2 \langle X^2 \rangle + \dots \right]^n \\
 &= \exp \left\{ n \ln \left[1 + \frac{2i\pi S_n}{n} \langle X \rangle - \frac{1}{2} \left(\frac{2\pi S_n}{n} \right)^2 \langle X^2 \rangle + \dots \right] \right\}
 \end{aligned} \tag{3.5}$$

Sabendo que para $z \ll 1$ logo,

$$\ln(1+z) = z - \frac{1}{2}z^2 + \frac{1}{3}z^3 + \dots \quad (3.6)$$

nesse caso da equação 3.5 obtem-se que

$$z = \frac{2i\pi S_n}{n} \langle X \rangle - \frac{1}{2} \left(\frac{2\pi S_n}{n} \right)^2 \langle X^2 \rangle + \dots,$$

isso porque n é muito grande e tende ao infinito. Desprezando os termos de terceira ordem em diante:

$$\begin{aligned} \ln \left(1 + \frac{2i\pi S_n}{n} \langle X \rangle - \frac{1}{2} \left(\frac{2\pi S_n}{n} \right)^2 \langle X^2 \rangle \right) &= \frac{2i\pi S_n}{n} \langle X \rangle - \frac{1}{2} \left(\frac{2\pi S_n}{n} \right)^2 \langle X^2 \rangle - \\ &\quad - \frac{1}{2} \left(\frac{2i\pi S_n}{n} \langle X \rangle - \frac{1}{2} \left(\frac{2\pi S_n}{n} \right)^2 \langle X^2 \rangle \right)^2 \\ &= \frac{2i\pi S_n}{n} \langle X \rangle - \frac{1}{2} \left(\frac{2\pi S_n}{n} \right)^2 \langle X^2 \rangle + \frac{1}{2} \left(\frac{2\pi S_n}{n} \right)^2 \langle X \rangle^2 \end{aligned}$$

logo:

$$\ln \left(1 + \frac{2i\pi S_n}{n} \langle X \rangle - \frac{1}{2} \left(\frac{2\pi S_n}{n} \right)^2 \langle X^2 \rangle \right) = \frac{2i\pi S_n}{n} \langle X \rangle + \frac{1}{2} \left(\frac{2\pi S_n}{n} \right)^2 (\langle X \rangle^2 - \langle X^2 \rangle) \quad (3.7)$$

Aplicando a equação 3.7 no obtido na equação 3.5:

$$\begin{aligned} F^{-1}[P_X(S_n)] &= \exp \left\{ n \left[\frac{2i\pi S_n}{n} \langle X \rangle + \frac{1}{2} \left(\frac{2\pi S_n}{n} \right)^2 (\langle X \rangle^2 - \langle X^2 \rangle) \right] \right\} \\ &= \exp \left[2i\pi S_n \langle X \rangle + \frac{1}{2} \frac{(2\pi S_n)^2}{n} (\langle X \rangle^2 - \langle X^2 \rangle) \right] \\ &= \exp \left[2i\pi S_n \langle X \rangle - \frac{1}{2} \frac{(2\pi S_n)^2}{n} \sigma_X^2 \right] \end{aligned} \quad (3.8)$$

Tomando agora a transformada de Fourier, encontra-se a função original, pois a transformada de Fourier da transformada inversa de Fourier é a função em questão. Logo:

$$\begin{aligned} P_X(S_n) &= \int_{-\infty}^{\infty} e^{-2i\pi S_n X} F^{-1}[P_X(S_n)] dS_n \\ &= \int_{-\infty}^{\infty} e^{-2i\pi S_n X} e^{2i\pi S_n \langle X \rangle - \frac{1}{2} \frac{(2\pi S_n)^2}{n} \sigma_X^2} dS_n \\ &= \int_{-\infty}^{\infty} e^{2i\pi S_n (\langle X \rangle - X) - \frac{1}{2} \frac{(2\pi S_n)^2}{n} \sigma_X^2} dS_n \end{aligned}$$

que é uma integral gaussiana do tipo:

$$\int_{-\infty}^{\infty} e^{iay-by^2} dy = e^{-a^2/4b} \sqrt{\frac{\pi}{b}} \quad (3.9)$$

onde para usualmente é escrito $a = 2\pi(\langle X \rangle - X)$ e $b = (2\pi\sigma_X)^2/2n$, logo,

$$\begin{aligned} P_X(S_n) &= \sqrt{\frac{\pi}{\frac{(2\pi\sigma_X)^2}{2n}}} \exp \left\{ \frac{-[2\pi(\langle X \rangle - X)]^2}{4(2\pi\sigma_X)^2/2n} \right\} \\ &= \sqrt{\frac{n}{2\pi\sigma_X^2}} e^{-\frac{n(X-\langle X \rangle)^2}{2\sigma_X^2}} \\ &= \sqrt{\frac{n}{2\pi\sigma_X^2}} e^{-\frac{nS_n^2}{2\sigma_X^2}} \end{aligned} \quad (3.10)$$

Que encerra a prova formalizada do teorema.

3.1.1 Alguns exemplos de convergências via TLC

Na seção 2.5.2 foi mostrado como a distribuição Gaussiana é dada no limite quando n é muito grande na distribuição binomial, isso em si é um exemplo de aplicação do TLC. Agora com a prova formalizada, pode-se observar como o teorema é visto em várias combinações de variáveis aleatórias que sejam i.i.d(independentes e igualmente distribuídas).

O primeiro exemplo é mostrar que o teorema é válido quando são usadas variáveis aleatórias distribuídas segundo uma gaussiana. Seja $S_n = X_1 + X_2 + \dots + X_n$ e Z uma variável que seja a soma de duas variáveis dentro do conjunto X , logo:

$$Z = X_k + X_l, \quad (3.11)$$

então, a probabilidade conjunta de X_k e X_l é:

$$P(X_l, X_k) = P(X_l)P(X_k) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-[X_l-\langle X \rangle]^2/2\sigma_X^2} \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-[X_k-\langle X \rangle]^2/2\sigma_X^2}, \quad (3.12)$$

ou seja,

$$P(X_l, X_k) = \frac{1}{2\pi\sigma_X^2} e^{-([X_l-\langle X \rangle]^2 + [X_k-\langle X \rangle]^2)/2\sigma_X^2} \quad (3.13)$$

A distribuição para Z será:

$$P(Z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-[X_l-\langle X \rangle]^2/2\sigma_X^2} \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-[Z-X_l-\langle X \rangle]^2/2\sigma_X^2} dX_l, \quad (3.14)$$

portanto:

$$P(Z) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_X^2} \exp \left\{ -\frac{[X_l - \langle X \rangle]^2 + [Z - X_l - \langle X \rangle]^2}{2\sigma_X^2} \right\} dX_l \quad (3.15)$$

expandindo os produtos notáveis:

$$\begin{aligned}
[X_l - \langle X \rangle]^2 + [Z - X_l - \langle X \rangle]^2 &= X_l^2 - 2X_l\langle X \rangle + \langle X \rangle^2 + (Z - X_l)^2 \\
&- 2(Z - X_l)\langle X \rangle + \langle X \rangle^2 \\
&= X_l^2 - 2X_l\langle X \rangle + \langle X \rangle^2 + Z^2 - 2ZX_l + X_l^2 - 2Z\langle X \rangle \\
&+ 2X_l\langle X \rangle + \langle X \rangle^2 \\
&= 2X_l + 2\langle X \rangle^2 + Z^2 - 2ZX_l - 2Z\langle X \rangle \\
&= 2X_l + 2\langle X \rangle^2 + Z^2 - 2ZX_l - 2Z\langle X \rangle + \frac{Z^2}{2} - \frac{Z^2}{2} \\
&= 2X_l - 2ZX_l + \frac{Z^2}{2} + \langle X \rangle^2 - 2Z\langle X \rangle + \frac{Z^2}{2} \\
&= \left(\sqrt{2}X_l - \frac{Z}{\sqrt{2}} \right)^2 + \left(\frac{Z}{\sqrt{2}} - \sqrt{2}\langle X \rangle \right)^2 \\
&= 2 \left(X_l - \frac{Z}{2} \right)^2 - \frac{1}{2} (Z - 2\langle X \rangle)^2
\end{aligned} \tag{3.16}$$

Para a probabilidade então:

$$P(Z) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_X^2} e^{-\frac{(x_l - \frac{Z}{2})^2}{\sigma_X^2}} e^{-\frac{(Z - 2\langle X \rangle)^2}{4\sigma_X^2}} dX_l \tag{3.17}$$

Isolando a integral gaussiana:

$$P(Z) = \frac{e^{-\frac{(Z - 2\langle X \rangle)^2}{4\sigma_X^2}}}{2\pi\sigma_X^2} \int_{-\infty}^{\infty} e^{-\frac{(x_l - \frac{Z}{2})^2}{\sigma_X^2}} dX_l \tag{3.18}$$

Resolvendo a integral gaussiana, o resultado será:

$$\begin{aligned}
P(Z) &= \frac{e^{-\frac{(Z - 2\langle X \rangle)^2}{4\sigma_X^2}}}{2\pi\sigma_X^2} \sqrt{\sigma_X^2 \pi} \\
&= \frac{1}{\sqrt{4\pi\sigma_X^2}} e^{-\frac{(Z - 2\langle X \rangle)^2}{4\sigma_X^2}}
\end{aligned} \tag{3.19}$$

Ou seja, a soma de duas variáveis aleatórias independentes e distribuídas pela mesma distribuição gaussiana, é uma outra distribuição gaussiana, então caso sejam somadas N variáveis, como afirma o TLC, obtêm-se também uma distribuição gaussiana.

Essa característica de distribuições Gaussianas em estatística é um conceito muito importante, o da estabilidade de uma distribuição. Uma distribuição é dita estável se:

$$S_2 = X_1 + X_2, \tag{3.20}$$

e S_2 , X_1 e X_2 seguem a mesma distribuição estatística. Como mostrado nesta seção, a distribuição gaussiana segue exatamente esses parâmetros.

Um outro exemplo canônico de uma distribuição bem conhecida que também converge para a gaussiana com n grande, é a distribuição de Poisson.

Considere a seguinte distribuição de Poisson definida por:

$$P(n) = \frac{\mu^n}{n!} e^{-\mu} \quad (3.21)$$

Tomando o logaritmo natural em ambos os termos da equação observa-se:

$$\ln P(n) = -\mu + n \ln \mu - \ln n! \quad (3.22)$$

Onde pela fórmula de Stirling com $n \rightarrow \infty$,

$$\begin{aligned} \ln n! &= n \ln n - n = \ln \left[\sqrt{2n\pi} \left(\frac{n}{\exp} \right)^n \right] \\ &= n \ln n - n + \ln(\sqrt{2n\pi}) \end{aligned} \quad (3.23)$$

Logo:

$$\ln P(n) = -\mu + n \ln \mu - n \ln n + n - \ln(\sqrt{2n\pi}) \quad (3.24)$$

Pode-se definir o desvio da média como sendo

$$z = n - \mu,$$

deixando o valor médio tender a infinito e tratando z/μ como muito pequeno, mas z^2/μ finito. Com essas definições:

$$\begin{aligned} \ln P(n) &= -\mu + (\mu + z) \ln \mu - (\mu + z) \ln(\mu + z) + \mu + z - \ln \sqrt{2\pi(\mu + z)} \\ &= (\mu + z) \ln \mu - (\mu + z) \ln(\mu + z) + z - \ln \sqrt{2\pi(\mu + z)} \\ &= (\mu + z) \ln \left(1 - \frac{z}{\mu + z} \right) + z - \ln \sqrt{2\pi\mu} \\ &= (\mu + z) \left(-\frac{z}{\mu + z} + -\frac{z^2}{2(\mu + z)^2} \right) + z - \ln \sqrt{2\pi\mu} \\ &\sim -\frac{z^2}{2\mu} - \ln \sqrt{2\pi\mu}, \end{aligned} \quad (3.25)$$

usando a expansão do logaritmo natural e o fato de que $z \ll \mu$. Isolando $P(n)$:

$$P(n) = \frac{1}{\sqrt{2\pi\mu}} e^{-z^2/2\mu}, \quad (3.26)$$

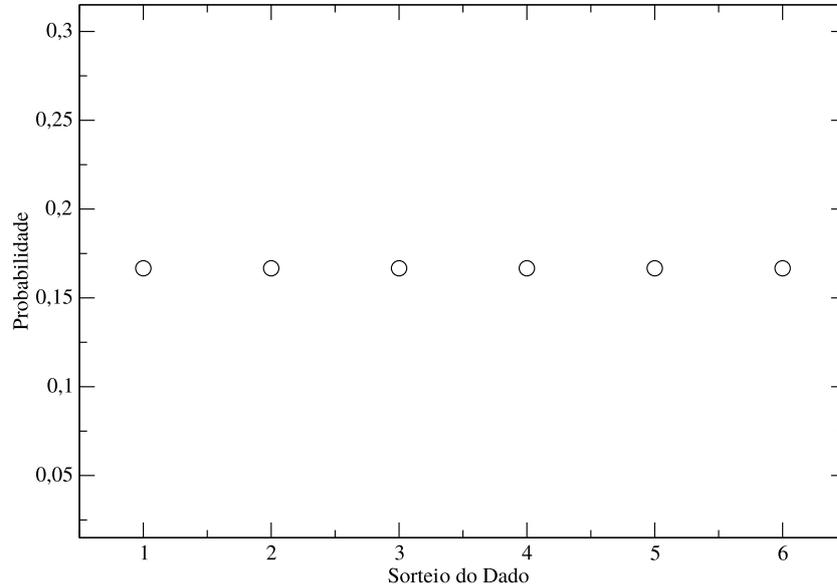
exatamente uma distribuição gaussiana, como afirmado anteriormente.

Por fim pode-se observar o TLC de uma maneira bem didática e pouco abordada por livros texto, o simples lançar de dados não viciados. Considere um dado de 6 faces marcadas com números de 1 até 6. Ao lançar esse dado uma única vez são percebidas 6 possibilidades iguais de ter alguma das faces, por ele não ser viciado. Por tanto os pontos mostram a distribuição uniforme de probabilidade, de mesma possibilidade de ocorrência de todos os valores possíveis do espaço amostral, na figura 3.1.

A medida que são acrescentado novos dados justos, visualmente é claro perceber que a curva começa a tomar uma forma bem característica com a gaussiana. A figura 3.2 fornece a noção dessa escala.

Figura 3.1: Distribuição discreta de probabilidades para o lançamento de um dado justo. Todos os valores possíveis de se obter, os números naturais de 1 à 6, tem a mesma probabilidade de ocorrência, $1/6$.

Probabilidades para 1 Dado



Fonte: Autor, 2014.

Os cálculos de convergência são similares aos mostrados para a distribuição binomial, mas, um pouco mais complexos, pois se tratam de 6 possibilidades a cada “passo”. Para não tornar o trabalho repetitivo e mostrar uma convergência já discutida, a probabilidade para n jogadas é

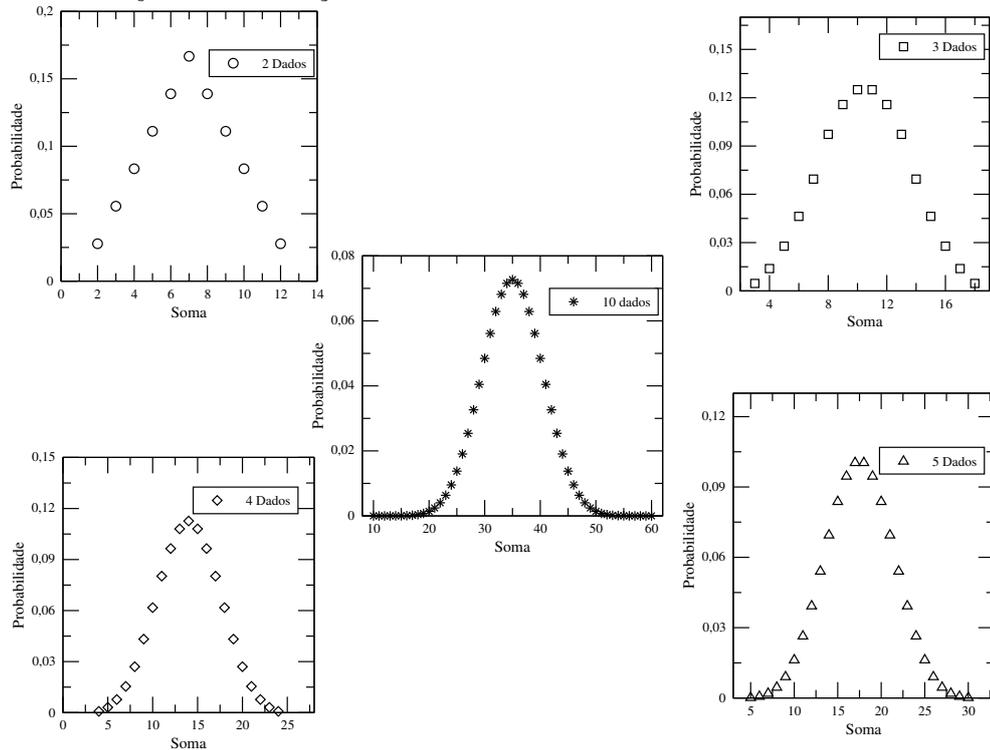
$$P_n(x) = \frac{1}{6^n} \sum_{i=0}^{\lfloor \frac{x-n}{6} \rfloor} (-1)^i \binom{n}{i} \binom{x-6i-1}{n-1}, \quad (3.27)$$

onde x é o número sorteado no dado, e a função $\lfloor k \rfloor$, escolhe o maior inteiro menor que k , é a chamada função chão. Usando a fórmula de Stirling para fatoriais quando $n \rightarrow \infty$, é possível mostrar a convergência.

3.2 As seções de Lévy

Em 1935 Paul Lévy publicou um trabalho onde ele apresenta um teorema chamado Teorema das Seções de Lévy[37], tal teorema é baseado em uma proposta de generalização do TLC. Nesse enfoque, Lévy mostrou como variáveis aleatórias que não eram necessariamente independentes, convergiam também para uma distribuição gaussiana, levando em conta suas variâncias. Esse é o ponto crucial da análise de Lévy, pois o TLC exige que as variáveis sejam i.i.d. enquanto que na abordagem proposta é possível encontrar a convergência em sistemas correlacionados.

Figura 3.2: Curvas representando as diversas distribuições para a soma das faces voltadas para cima em conjuntos de lançamentos de dados



Fonte: Autor, 2014.

3.2.1 O Teorema das Seções de Lévy(TSL)

O teorema das seções de Lévy, proposto por Paul Lévy[38], conduz a uma maneira mais abrangente de se tratar a convergência de sistemas para o regime gaussiano. Usualmente esse teorema é muito confundido, pelo leitor não especialista, com as distribuições de Lévy e suas inúmeras aplicações, na verdade o TSL, vem tratar, assim como o TLC, de como variáveis aleatórias convergem para distribuições normais de probabilidade, com a diferença que esse teorema contempla distribuições com variância infinita e variáveis aleatórias correlacionadas.

Para compreender a ação do teorema, considere uma série, ou uma cadeia de variáveis aleatórias dada por:

$$r = (X_1, X_2, \dots, X_{i-1}, X_i, \dots, X_{k-1}, X_k, \dots, X_j, \dots, X_n) \quad (3.28)$$

e com médias locais dadas por:

$$\mu_i = \int_{-\infty}^{\infty} x_i P(x_i | x_1, x_2, \dots, x_{i-1}) dx_i. \quad (3.29)$$

onde $P(x_i | x_1, x_2, \dots, x_{i-1})$ é a probabilidade condicional de obter uma realização x_i de X_i , ocorrido as realizações $x_1 \dots x_{i-1}$.

Dessa forma, também é definida uma variância local para a série:

$$\begin{aligned}
 m_i^2 &= \langle X_i^2 \rangle - \langle X_i \rangle^2 \\
 &= \int_{-\infty}^{\infty} x_i^2 P(x_i|x_1, x_2, \dots, x_{i-1}) dx_i - \left(\int_{-\infty}^{\infty} x_i P(x_i|x_1, x_2, \dots, x_{i-1}) dx_i \right)^2 \\
 &= \int_{-\infty}^{\infty} x_i^2 P(x_i|x_1, x_2, \dots, x_{i-1}) dx_i - \mu_i^2.
 \end{aligned} \tag{3.30}$$

Com isso, é possível escrever uma série m' com todas as variâncias locais da série r :

$$m' = (m_1^2, m_2^2, \dots, m_{i-1}^2, m_i^2, \dots, m_{k-1}^2, m_k^2, \dots, m_j^2, \dots, m_n^2). \tag{3.31}$$

É definido então uma quantidade λ_i dada pela expressão:

$$\lambda_i = \sum_{l=1}^i m_l^2 \tag{3.32}$$

Ou seja, somando todos as variâncias locais das i realizações anteriores a x_i .

Um outro parâmetro também deve ser apresentado, o tamanho t da seção, definido como sendo:

$$\lambda_{i-1} \leq t < \lambda_i, \tag{3.33}$$

ou seja, um valor máximo para a somatória das variâncias agregadas. Dessa forma o conjunto de realizações compreendida de x_1 à x_i é um conjunto pertencente a **seção t** que tem uma definição que parte da condição da seção t representada pela equação 3.33.

Dada essa condição, é definida uma nova variável aleatória S_{1_t} na forma:

$$S_{1_t} = \sum_{l=1}^i X_l. \tag{3.34}$$

Em seguida procura-se um $k > i$, de tal forma que:

$$\lambda_{k-1} \leq t < \lambda_k, \tag{3.35}$$

e é calculado mais uma variável aleatória, agora chamada S_{2_t} , dada por:

$$S_{2_t} = \sum_{l=i+1}^k X_l. \tag{3.36}$$

E o procedimento é repetido até ser obtido λ_n que é definido pela condição dada na equação 3.33 e calculada a variável S_{w_t} , dada por:

$$S_{w_t} = \sum_{l=j+1}^n X_l. \tag{3.37}$$

Com todas as w variáveis aleatórias definidas pela seção t , é fixada uma nova série da seção t dada por:

$$S_t = (S_{1_t}, S_{2_t}, \dots, S_{w_t}). \tag{3.38}$$

O teorema das seções de Lévy é anunciado, com respeito a série S_t , como[38]:

Teorema 2 (TSL) *Para uma média local nula, ou seja, $\mu_n = 0$ e para variáveis aleatórias que satisfaçam a condição de Lindeberg (apêndice A), a distribuição para a variável $\frac{S_t}{\sqrt{t}}$ é dada por:*

$$\lim_{t \rightarrow \infty} P\left(\frac{S_t}{\sqrt{t}} < \eta\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\eta} e^{-\frac{x^2}{2}} dx \quad (3.39)$$

Sendo assim o teorema estende a sua aplicação para até variáveis correlacionadas, pois não há necessidade de variáveis i.i.d., algo não permitido no TLC em sua forma clássica.

Em termos didáticos, é possível construir um exemplo lúdico para deixar clara a aplicação do teorema. Seja r' uma série de variáveis aleatórias arbitrariamente escolhidas dada por:

$$r' = (X_1, X_2, X_3, X_4, X_5, X_6, \dots, X_n). \quad (3.40)$$

Definido o tamanho t , considere que:

$$\begin{aligned} m_1^2 &< t \text{ e que} \\ m_1^2 + m_2^2 &< t \text{ mas,} \\ m_1^2 + m_2^2 + m_3^2 &\geq t \end{aligned}$$

logo,

$$S_{1_t} = X_1 + X_2 + X_3 \quad (3.41)$$

Aplicando a regra novamente, caso tenha:

$$\begin{aligned} m_4^2 &< t \text{ mas,} \\ m_4^2 + m_5^2 &\geq t \end{aligned}$$

logo o segundo termo será,

$$S_{2_t} = X_4 + X_5. \quad (3.42)$$

Continuando com a metodologia, seja:

$$\begin{aligned} m_6^2 &< t \text{ e que} \\ m_6^2 + m_7^2 &< t \text{ e,} \\ m_6^2 + m_7^2 + m_8^2 &< t \text{ mas,} \\ m_6^2 + m_7^2 + m_8^2 + m_9^2 &\geq t \end{aligned}$$

o próximo termo da série é dado por:

$$S_{3_t} = X_6 + X_7 + X_8 + X_9. \quad (3.43)$$

A aplicação da metodologia é continuada até X_n para o cálculo do termo S_{w_t} .

Afim de compreender como se dá a ação do teorema das seções de Lévy, segue a prova apresentada pelo próprio Lévy em 1935[38].

Prova do Teorema Considere uma sequência de variáveis aleatórias com a propriedade:

$$E(X_1) = 0 \quad (3.44)$$

e

$$E(X_\nu | X_1, \dots, X_{\nu-1}) = 0 \quad (3.45)$$

com $\nu = 2, 3, \dots$. Ao invés de considerar a soma usual das variáveis X_ν , Lévy considerou uma soma do tipo:

$$S(t) = \sum_{\nu=1}^{N(t)} X_\nu, \quad (3.46)$$

onde $N(t)$ é um número aleatório definido por:

$$N(t) = \min \{n \in \mathbb{N} | m_1^2 + m_2^2 + \dots + m_n^2 \geq t\} \quad (3.47)$$

e t é um número positivo que Lévy interpretou como sendo o tempo necessário para a convergência de um certo número de tentativas.

Dessa forma, pode-se definir uma nova variável aleatória, X'_ν , de tal forma que:

$$X'_\nu = \begin{cases} \frac{X_\nu}{\sqrt{t}} & \text{se } \nu \leq N(t) \\ 0 & \text{caso contrário} \end{cases} \quad (3.48)$$

Sendo assim a variância será:

$$m_\nu'^2 = \begin{cases} \frac{m_\nu^2}{t} & \text{se } \nu \leq N(t) \\ 0 & \text{caso contrário} \end{cases} \quad (3.49)$$

De posse dessas novas variáveis:

$$\sum_{k=1}^n X'_k + \sum_{j=n+1}^{\infty} m'_j \xi_j + Z = S'_n + R'_n + Z, \quad (3.50)$$

onde ξ_j é uma variável aleatória distribuída gaussianamente e Z obedece uma distribuição F que deve ser bem comportada e com derivadas de até terceira ordem finitas. É importante considerar que as variáveis ξ_j , Z e X_ν , são mutualmente independentes, o que não ocasiona perda de generalidade. Logo:

$$|P(S'_n + R'_n + Z < \eta) - P(S'_{n-1} + R'_{n-1} + Z < \eta)| \leq \text{MAX}\{X'_n\} m'_n 2, \quad (3.51)$$

com $\eta' = S'_n + R'_n + Z$, e usando a convolução, já antes mostrada, calcula-se $P(\eta' < \eta)$.

Dessa maneira, voltando as variáveis originais:

$$\left| P\left(\sum_{\nu=1}^n X'_\nu + Z < \eta\right) - P\left(\sum_{\nu=1}^n m'_\nu \xi_\nu + Z < \eta\right) \right| \leq \text{MAX}\{X'_n\} \sum_{\nu=1}^n m'_\nu 2, \quad (3.52)$$

com o cancelamento de termos repetidos.

Tomando o limite quando n tende ao infinito, com as propriedades das equações 3.48 e 3.49, logo

$$\left| P\left(\sum_{\nu=1}^{N(t)} X'_\nu + Z < \eta\right) - P\left(\sum_{\nu=1}^{N(t)} m'_\nu \xi_\nu + Z < \eta\right) \right| \leq \text{MAX}\{X'_n\} \sum_{\nu=1}^{N(t)} m'_\nu 2 \quad (3.53)$$

trocando as variáveis com linha para variáveis sem linha:

$$\left| P \left(\sum_{\nu=1}^{N(t)} \frac{X_{\nu}}{\sqrt{t}} + Z < \eta \right) - P \left(\sum_{\nu=1}^{N(t)} \frac{m_{\nu} \xi_{\nu}}{\sqrt{t}} + Z < \eta \right) \right| \leq \frac{\text{MAX}\{X_n\}}{\sqrt{t}} \sum_{\nu=1}^{N(t)} \frac{m_{\nu}^2}{t} \quad (3.54)$$

Por definição a soma das variáveis $X_{\nu}u$, é $S(t)$ e pode-se tomar como $\sum_{\nu=1}^{N(t)} m_{\nu}^2 = t$, pela equação 3.47. Sendo assim:

$$\left| P \left(\frac{S(t)}{\sqrt{t}} + Z < \eta \right) - P \left(\sum_{\nu=1}^{N(t)} \frac{m_{\nu} \xi_{\nu}}{\sqrt{t}} + Z < \eta \right) \right| \leq \frac{\text{MAX}\{X_n\}}{\sqrt{t}} \quad (3.55)$$

Tomando t para infinito, o termo da desigualdade se anula e temos finalmente que:

$$\lim_{t \rightarrow \infty} P \left(\frac{S(t)}{\sqrt{t}} + Z < \eta \right) = P(\Xi + Z < \eta), \quad (3.56)$$

Como Z é uma função arbitrária que pode ser escolhida como pequena e Ξ segue uma distribuição gaussiana, assim se dá a prova.

3.2.2 TSL aplicado à séries temporais

O TSL foi mostrado para o caso de variáveis contínuas, mas este trabalho está focado em estudar séries temporais, ou seja, conjuntos discretos de variáveis estocásticas, sendo assim, é necessário re-escrever o TSL para o caso de séries de números.

Considerando uma série de N realizações da variável aleatória X_i , é preciso calcular a seção t dessa série temporal, para isso inicialmente calcula-se a variância local, o m_n . Defini-se então uma nova série[39, 40, 41] $\{x'_1, x'_2, \dots, x'_{n-2\eta}\}$ com uma relação entre nova série e a série antiga, na forma $x'_k = x_{k+\eta}$. Tal nova série tem 2η elementos a menos, retirando os primeiros e últimos η números. Dessa forma é possível definir agora a variância local como sendo:

$$m_{l,n}^2 = \frac{1}{2\eta+1} \sum_{i=1}^{2\eta+1} x_{l+n-1+i}^2 - \left(\frac{1}{2\eta+1} \sum_{i=1}^{2\eta+1} x_{l+n-1+i} \right)^2, \quad (3.57)$$

onde $l+n$ varre os valores de 1 até $n-2\eta$ e $l = \sum_i (-1)^k n_i$, com n_i representando a quantidade de números usados para obter o i -ésimo termo da seção, ou seja, o índice l ajuda a garantir que não existirá sobreposição de termos na formação da nova série.

Com a variância local calculada, e utilizando os conceitos das equações 3.32 e 3.33 para determinar a seção t , calcula-se o conjunto S_t dado por:

$$S_t = \left(\sum_{i=1}^{n_1} x'_i, \sum_{i=1}^{n_2} x'_{1+i}, \dots, \sum_{i=1}^{n_k} x'_{k+i} \right) \quad (3.58)$$

E esse série tem uma distribuição de probabilidades que tende ao regime gaussiano quando t tende ao infinito, assegurado pelo TSL.

3.3 O Teorema do Limite Central Clássico vs Teorema das Seções de Lévy

O Teorema das Seções de Lévy primeiramente deve ser encarado como uma generalização do conceito proposto no TLC, ele serve como uma nova ferramenta para analisar a convergência de qualquer série e principalmente, como é o presente caso, séries não estacionárias e fortemente correlacionadas. A principal diferença entre os métodos é a sua forma de agregação, no TLC em sua maneira clássica nós agregamos as variáveis uma a uma gerando uma nova série S_n , que pode ser escrita como

$$S_n = x_1 + x_2 + \dots + x_n - \langle x \rangle, \quad (3.59)$$

como já mostrado na equação 3.1.

Como também já foi discutido, essa metodologia de agregação não garante que a série S_n venha convergir para uma gaussiana se a mesma for composta de variáveis x_i correlacionadas. Para garantir o TLC as variáveis devem ser i.i.d, que são variáveis independentes e assim determina-se facilmente as integrais mostradas na prova do teorema.

Em contra partida temos o TSL onde as variáveis são agregadas dependendo da variância local e do parâmetro t da seção, formando assim a série S_t que é escrita como:

$$S_t = \left(\sum_{i=1}^{n_1} x'_i, \sum_{i=1}^{n_2} x'_{1+i}, \dots, \sum_{i=1}^{n_k} x'_{k+i} \right) \quad (3.60)$$

que como agrega as variáveis por sua variância, os termos pertencentes a série S_t são heterogêneos em número de variáveis isso porque é preciso definir um $\lambda_{l,n}$ como sendo:

$$\lambda_{l,n} = \sum_{i=1}^n m_{i,l}^2, \quad (3.61)$$

e o $\lambda_{l,n}$ está condicionado ao número t . Então em uma série, para satisfazer a condição t talvez sejam agregadas 3 variáveis, depois para o próximo λ é possível que sejam agregados 5 números e no caso posterior sejam agregadas agora apenas 2 variáveis e assim por diante, não tendo um número definido de variáveis que irão compor os termos que formam a série S_t .

A grande vantagem nesse modelo é o fato de poder usar quaisquer tipos de variáveis, tanto correlacionadas quanto não-estacionárias, tornando a convergência gaussiana muito mais provável e abrangente. A discussão comparativa é extremamente importante para a construção da compreensão sobre a convergência dos dados e por isso aqui mostramos um confronto mais claro sobre as diferenças entre os Teoremas.

3.4 Testes de Convergência

Como já foi mostrado nas seções anteriores, os dois teoremas, tanto o TLC, quando o TSL, apresentam a convergência para o regime gaussiano quando utilizam-se infinitas variáveis envolvidas. Em séries temporais finitas, tais como, preços de ativos, corrente em alguns circuitos e outros, onde existe um tempo característico de medida, se faz necessário testar se já foi atingido ou não o regime gaussiano, na aplicação de ambos os teoremas.

Nessa perspectiva, é importante que a nova série seja comparada com a que é obtida a cada iteração do teorema, com os parâmetros bem conhecidos na literatura, para as séries gaussianas. Aqui serão apresentados alguns dos principais testes comparativos que podemos utilizar, afim de garantir a convergência para o regime desejado.

3.4.1 Momentos estatísticos

Os primeiros testes são a verificação dos momentos estatísticos. Os momentos de primeira e segunda ordem são quantidades que normalmente não são usados como parâmetros comparativos, já que cada curva gaussiana tem uma média e um desvio padrão individual. Para comparação são utilizados alguns momentos de mais alta ordem, como por exemplo, a *skewness* e a *kurtosis*, que são o terceiro e quarto momentos respectivamente.

Como discutido em seções anteriores, a distribuição gaussiana, independente de sua origem, possui alguns parâmetros que são imutáveis, todos os seus momentos ímpares são sempre nulos, isso se dá pelo fato que a distribuição gaussiana é uma distribuição simétrica e os momentos ímpares calculam diferentes graus de assimetria. Sendo assim, se uma distribuição é Gaussiana ela deverá ter sua *skewness* sempre nula.

Outro momento estatístico importante é a *kurtosis*, ou o quarto momento. Esse parâmetro mede a “curvatura” da distribuição. A equação 2.30 com dados gaussianamente distribuídos, a *kurtosis* é 3, por essa razão alguns autores preferem definir o quarto momento subtraindo de 3 a equação 2.30, gerando uma “kurtose excessiva”, definindo assim a kurtose gaussiana como sendo nula também, como discutido no capítulo anterior.

3.4.2 Expoente de Hürst e a Fractalidade

O expoente de Hürst é muito importante para a análise dimensional de uma figura fractal. Em uma série temporal o expoente de Hürst ajuda a entender o comportamento que a série tem e como seus parâmetros estão relacionados. Se $H < \frac{1}{2}$, a série é anti-persistente, ou seja, tende a fazer o oposto ao que fez no passado, então a tendência da série segue sempre contrária. Com $H > \frac{1}{2}$ a série é persistente, tende a repetir ações passadas, ou seja, mantém a mesma tendência. Nesses dois exemplos as series são correlacionadas e esse grau de correlação pode ser inferido pelo estudo do expoente de Hürst.

Para o caso gaussiano, pelo próprio TLC, sabe-se que uma distribuição gaussiana é formada por variáveis independentes, ou seja, decorrelacionadas. O valor do expoente de Hürst para esse tipo de distribuição é $H = \frac{1}{2}$, onde todas as correlações são nulas e o TLC é facilmente aplicado. No estudo de convergência, o expoente de Hürst mostra como a série se comporta com a agregação de variáveis, fornecendo mais um índice capaz de informar se a série é ou não gaussiana.

Outro fator importante nessa análise é o grau de fractalidade do sistema. Como mostrado anteriormente, existem sistemas fractais que podem ser monofractais ou multifractais. Os cálculos e as análises feitas para determinar comportamento fractal foram apresentadas no capítulo anterior.

Variáveis distribuídas gaussianamente tem o expoente de Hürst igual a $1/2$ e como são variáveis decorrelacionadas, esse valor se mantém imutável independente do momento estatístico calculado. Ou seja, usando o método MF-DFA, obtêm-se o mesmo valor $1/2$ para a série em todas as análises de momentos.

Com essa constância do expoente na equação 2.139, existe uma dependência linear com o q -ésimo momento, tornando o espectro multifractal $f(\alpha)$ com uma curva muito estreita e centrada também no valor $1/2$. Com um espectro único valor para descrever a fractalidade, o sistema é monofractal e por tanto, séries distribuídas gaussianamente são séries monofractais.

Para analisar a convergência da série basta calcular a dependência o expoente de Hürst com o q -ésimo momento a cada tempo ou verificar o estreitamento da curva apresentada pelo espectro multifractal e determinar a transição entre o regime multifractal de fortes correlações e o regime monofractal descorrelacionado.

3.4.3 Teste KS

O teste Kolmogorov-Smirnov ou apenas KS, foi proposto nos anos 30 pelo russo Andrey Kolmogorov[42] e aprimorado por Nikolai Smirnov[43] em sua análise de distribuições. O teste KS é um teste não paramétrico ou seja, um teste que não se baseia em uma distribuição específica e é utilizado para comparar uma amostra de uma determinada série com uma distribuição, o chamado teste KS de uma amostra, ou também é usado para comparar duas amostras vindas de séries diferentes, o teste KS de duas amostras.

Para entender o teste segue o seguinte exemplo: considere um conjunto de realizações de uma variável aleatória X , que tem alguma distribuição desconhecida P e deseja-se testar a hipótese se essa distribuição desconhecida na verdade é igual a uma distribuição particular P_0 .

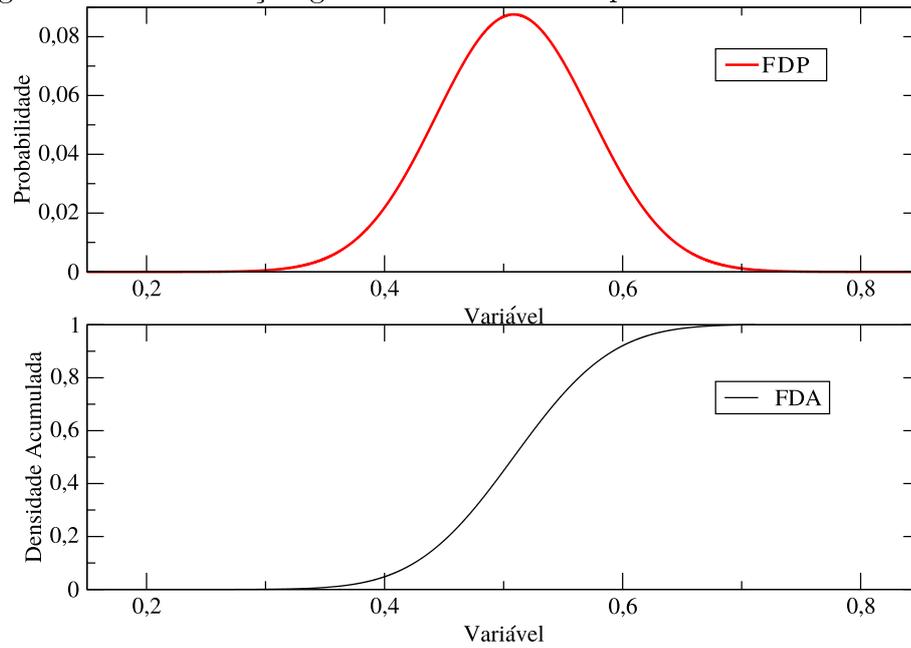
Primeiro calcula-se a função de distribuição acumulada empírica. Para conjuntos contínuos a FDA é a integral da função densidade de probabilidade, ou seja,

$$F(x) = \int_{-\infty}^x P(t)dt \quad (3.62)$$

Para uma única distribuição gaussiana a FDA pode ser vista na figura 3.3.

Verifica-se que a FDA segue o comportamento da distribuição, essa tendência fica clara ao observar várias distribuições gaussianas com diferentes parâmetros associados como mostrado na figura 3.4.

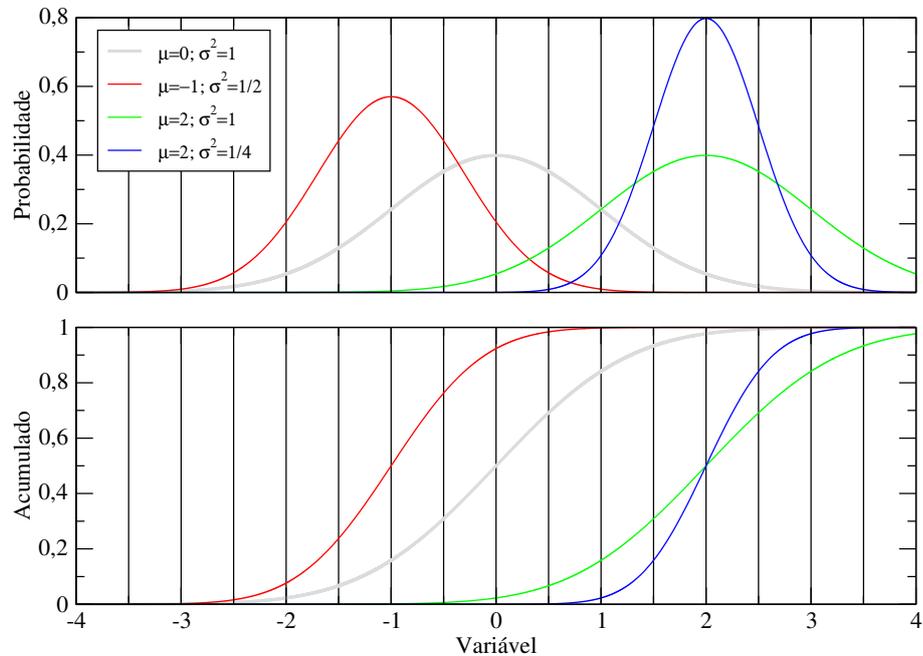
Figura 3.3: Distribuição gaussiana e sua FDA pra uma variável aleatória



Fonte: Autor, 2014.

Quando são tratadas as variáveis discretas, a FDA se torna agora a função de distri-

Figura 3.4: Diversas distribuições gaussianas com diferentes parâmetros e suas FDA afim de comparação e observação do comportamento tomado por cada uma das funções acumuladas.



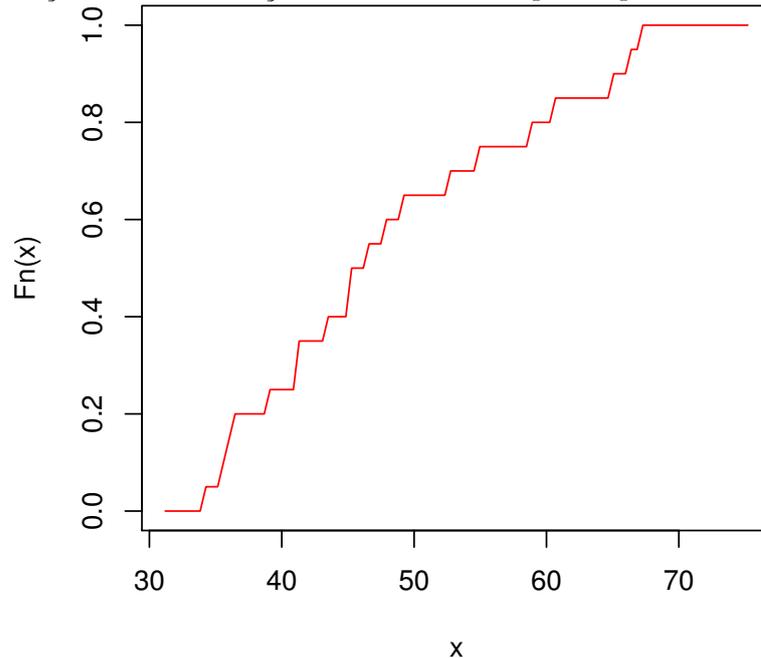
Fonte: Autor, 2014.

buição acumulada empírica(FDAE), com uma definição dada como:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x) \quad (3.63)$$

onde $I(X_i \leq x)$ é a função indicador que vale 1 para $X_i \leq x$ e é nula caso contrário. Essa função empírica apresenta uma curva que muito se assemelha a uma função do tipo degrau(figura 3.5).

Figura 3.5: A Função de Distribuição Acumulada Empírica para um conjunto de dados



Fonte: Autor, 2014.

O teste de Kolmogorov-Smirnov consiste em um teste comparativo entre as funções de densidade acumulada. O primeiro caso é o teste KS de uma amostra, onde uma FDAE é comparada com uma FDA de uma distribuição conhecida.

Se $F_n(x)$ é a FDAE de uma série desconhecida e $F(x)$ é a FDA de uma distribuição contínua conhecida com a qual deseja-se comparar à série empírica, define-se uma quantidade D_n como sendo:

$$D_n = \sup_x |F_n(x) - F(x)|, \quad (3.64)$$

onde a função \sup_x é a função *supremo*. Em termos matemáticos o supremo de um subconjunto S de um conjunto total parcialmente ordenado T , é o o menor elemento de T que seja maior que qualquer elemento de S . Esse conceito é muito importante na construção dos números reais e até mesmo na noção de limite em análise real.

Pelo teorema de Glivenko–Cantelli[44], para todo x fixo, se a amostra realmente é da distribuição $F(x)$, quase certamente

$$D_n = \sup_x |F_n(x) - F(x)| \rightarrow 0 \quad (3.65)$$

Ou seja caso o teste forneça $D_n = 0$ logo as duas distribuições são a mesma. Graficamente o teste pode ser visto como mostrado na figura 3.6

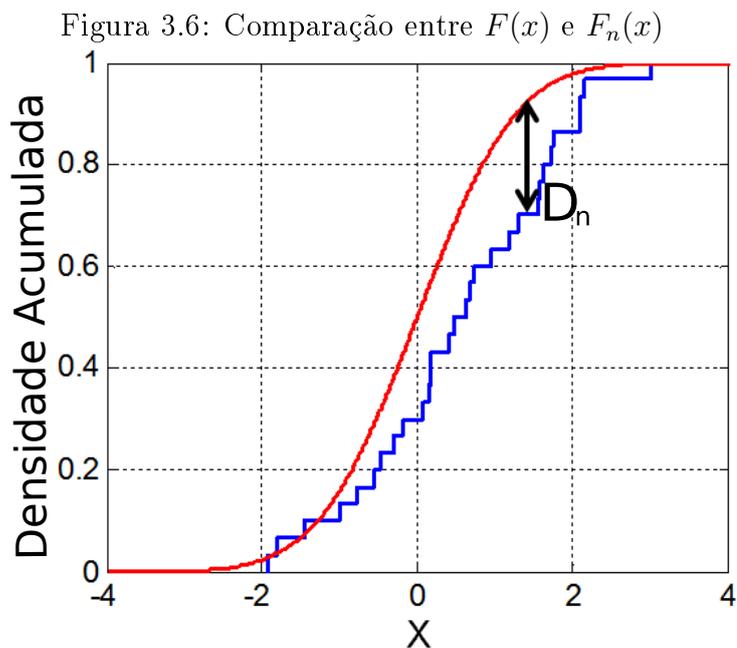
O Teste KS de duas amostras compara duas funções de distribuição acumulada empíricas entre si. Como o teste KS é um teste não paramétrico, ele possibilita esse tipo de análise comparativa entre duas FDAEs quais quer. Em termos matemáticos a construção do teste é muito similar ao teste para uma amostra,

$$D_{n,n'} = \sup_x |F_{1,n}(x) - F_{2,n'}(x)| \quad (3.66)$$

Onde são comparadas as duas FDAEs, $F_{1,n}(x)$ e $F_{2,n'}(x)$. Como descrito no teste de uma amostra, se as duas funções são da mesma distribuição, logo quase certamente $D_{n,n'} \rightarrow 0$.

Em termos gráficos o teste KS para duas amostras é representado na figura 3.7, onde claramente o parâmetro $D_{n,n'}$ fica descrito.

Para convergência o teste KS é um ótimo parâmetro, por ser um teste não-paramétrico ele se torna mais sensível as condições da FDAE. Para garantir se o sistema convergiu para o regime gaussiano, basta usar o teste para uma amostra, ou ainda gerar uma amostra distribuída gaussianamente e compara-la com a amostra desconhecida usando o teste KS de duas amostras. A medida que o parâmetro D_n vai se tornando mais próximo de zero, infere-se que a convergência está sendo alcançada.



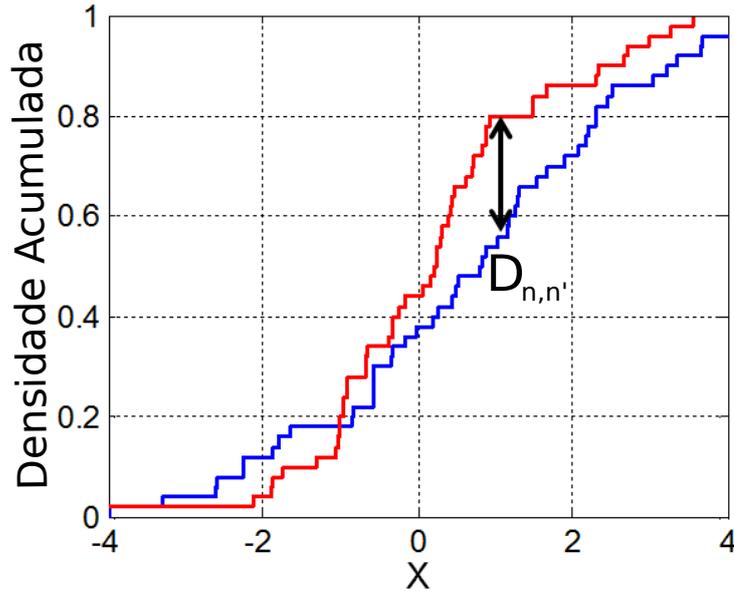


Figura 3.7: Comparação entre $F_{1,n}(x)$ e $F_{2,n'}(x)$

3.4.4 Teste KL de divergência

O teste KL de divergência foi proposto por Kullback e Leibler em 1951[45] e é uma medida não simétrica da diferença entre duas distribuições P e Q . O teste na verdade mede o quanto de informação é perdida quando Q é usado para se aproximar a P , é uma medida então da entropia de Shanon[46] do sistema.

Para variáveis discretas o teste apresenta o parâmetro $D_{\text{KL}}(P\|Q)$ para a divergência de Q com relação a P , como sendo:

$$D_{\text{KL}}(P\|Q) = \sum_i P(i) \ln \left(\frac{P(i)}{Q(i)} \right), \quad (3.67)$$

que pode ser escrito também como:

$$D_{\text{KL}}(P\|Q) = - \sum_i p_i \ln q(x) + \sum_i p_i \ln p_i \quad (3.68)$$

e para distribuições contínuas toma a forma:

$$D_{\text{KL}}(P\|Q) = \int_{-\infty}^{\infty} p(x) \ln \left(\frac{p(x)}{q(x)} \right) dx, \quad (3.69)$$

onde p e q são as densidades de probabilidade das séries P e Q .

O parâmetro $D_{\text{KL}}(P\|Q)$ é sempre positivo e só é igual a nulidade caso $P = Q$ em todos os pontos. Esse fator é comprovado pela desigualdade de Gibbs.

A desigualdade de Gibbs afirma que considerando dois conjuntos P e Q , temos que:

$$- \sum_{i=1}^n p_i \ln p_i \leq - \sum_{i=1}^n p_i \ln q_i \quad (3.70)$$

e só assumem a igualdade caso $P = Q$ para todos os elementos. Isso pode ser mostrado somando sobre um espaço onde $\sum_i p_i = 1$

$$\begin{aligned}
 -\sum_i p_i \ln \frac{q_i}{p_i} &\geq -\sum_i p_i \left(\frac{q_i}{p_i} - 1 \right) \\
 &= -\sum_i q_i + \sum_i p_i \\
 &= -\sum_i q_i + 1 \\
 -\sum_i p_i \ln \frac{q_i}{p_i} &\geq 0.
 \end{aligned} \tag{3.71}$$

Então

$$-\sum_i p_i \ln q_i \geq -\sum_i p_i \ln p_i,$$

o que formaliza o teste.

Esse teste é bastante útil no âmbito da convergência, por se tratar de um teste que analisa a quantidade de informação da série, como são tratadas séries temporais correlacionadas, sua convergência para o regime gaussiano faz com que algumas correlações sejam perdidas e uma medida para essa quebra de correlações é a medida da entropia de Shannon. Sendo assim o teste KL de divergência é um bom parâmetro comparativo de convergência.

4

A Econofísica como exemplo de séries correlacionadas

Em meados do final do século XIX, um economista social, Vilfredo Pareto, estimou que as riquezas em uma sociedade nos moldes dos países europeus dessa época, eram distribuídas de forma desigual e seguindo uma lei de potência, tal como:

$$y \sim x^{-\xi} \quad (4.1)$$

onde y é o número de pessoas com uma renda de x , ou superior a x e ξ é o expoente de Pareto, que foi estimado para a sociedade italiana, como sendo 1,5[9].

Em 1900 Bachelier escreveu em sua tese de doutorado uma análise estatística de ativos financeiros[47]. Nessa análise, ele propôs que o preço dos ativos financeiros deveriam seguir um comportamento tal qual, como é conhecido hoje, o movimento Browniano, de tal maneira que o mercado seria regido por uma distribuição gaussiana. Seu trabalho, apesar de algumas críticas atuais pela falta de rigor ao justificar a distribuição gaussiana como solução para a equação de Chapman-Kolmogorov[48], foi um trabalho pioneiro para a compreensão matemática do sistema financeiro.

A ideia de Bachelier foi contestada em 1963 com o trabalho revolucionário de Mandelbrot que sugeriu que os preços deveriam seguir um padrão de movimento chamado de vôos de Lévy, onde a distribuição resultante seria uma distribuição α -estável de Lévy[49]. Esse resultado abriu portas para novas interpretações mais acuradas acerca do comportamento do mercado.

Em 1973, Black e Scholes publicaram seu trabalho sobre opções financeiras[50]. Esse trabalho rendeu a eles o prêmio nobel em economia, fortalecendo ainda mais o estudo estatístico do comportamento de ativos em mercados financeiros, auxiliando na previsão e compreensão das reais dinâmicas da economia.

Nos anos 90, físicos começaram a voltar seus olhos para esse tipo de análise, com os parâmetros e as analogias físicas, como conceitos da teoria de escala, universalidade, sistemas desordenados, sistemas complexos e criticalidade auto-organizada encontrando motivação em trabalhos já publicados na área[51, 52, 53].

Em 2000, a área, conhecida hoje como *Econofísica*, teve seu primeiro livro publicado, por H.E. Stanley e R.N. Mantegna[54], fortalecendo ainda mais o campo. No Brasil, já foram realizados três encontros específicos da área, em 2007, 2010 e agora em 2014, organizado por diversos professores de várias instituições do país e contando com seminários

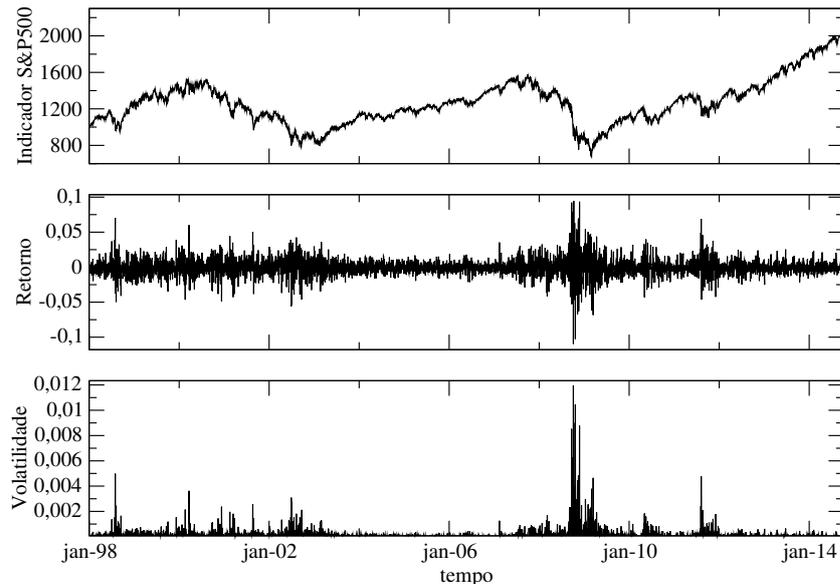
e palestras de físicos, economistas e inclusive agentes de bolsas de valores, agregando máximo possível de conhecimento aos pesquisadores mais antigos e os mais jovens[55, 56, 57].

Em suas análises, algumas quantidades relevantes ao sistema foram melhor definidas, como por exemplo o retorno financeiro. Retorno financeiro é definido como sendo:

$$r_{\Delta t}(t) = \log \left(\frac{S(t + \Delta t)}{S(t)} \right), \quad (4.2)$$

onde, $S(t)$ é o preço do dado ativo e Δt é o intervalo de tempo entre transações. Definido dessa forma, o retorno se mostra muito mais comportado que o próprio preço, e definido dessa forma ele serve como um parâmetro geral para todos os mercados, tendo em vista que é um parâmetro independente de moeda ou do valor agregado a um certo ativo financeiro, ele apenas compete como sendo uma variação de preços.

Além do retorno, outra quantidade é bastante utilizada e relevante ao estudo dos sistemas econômicos sob uma visão física, a volatilidade. Volatilidade é definida como sendo o desvio padrão da variação logarítmica dos preços, ou seja, do retorno. A volatilidade é relacionada com a quantidade de informação que é fornecida em um mercado em um dado tempo, por exemplo, ao receber uma grande quantidade de informação, o mercado agirá de tal forma com que o número médio de transações aumente proporcionalmente à quantidade de informação inserida.

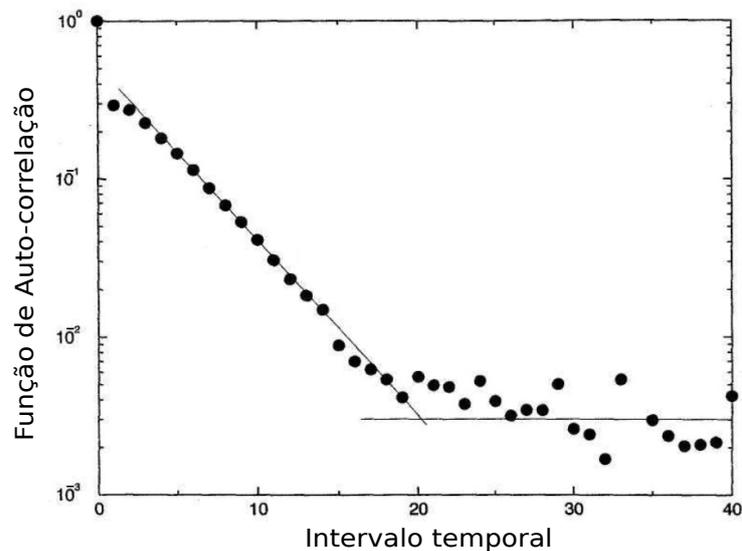


Fonte: Preços obtidos em www.yahoofinance.com acessado em: 12/05/2014, retorno e volatilidade calculados por Autor 2014.

Ao longo desses estudos algumas propriedades intrínsecas aos mercados financeiros foram observadas, de tal forma que qualquer mercado que seja estudado, teria todas as propriedades sugeridas. Essas propriedades são chamadas de Fatos Estilizados de Mercado. Os principais fatos estilizados são:

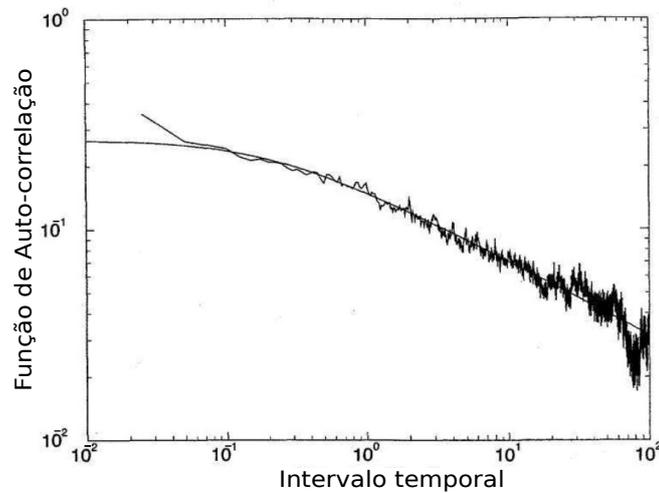
- Retornos de uma forma geral não apresentam correlação, ou seja, a autocorrelação da série temporal de retorno decai exponencialmente para zero (ver figura 4.1).
- Se hoje o mercado apresenta uma elevada volatilidade, então a probabilidade de observarmos um grande evento amanhã é maior que a média (Agregação de Volatilidade), informação gera mais informação. Esse fato também indica que existe correlação na série de volatilidade (ver figura 4.2).
- A distribuição de retornos de alta frequência apresenta caudas grossas, em geral são simétricas e leptocúrticas, enquanto a distribuição de retornos de baixa frequência tendem a distribuições com caudas mais exponenciais (ver figura 4.3).
- Os q-ésimos momentos da distribuição das flutuações dos preços (retornos) são multifractais (ver figura 4.4).

Figura 4.1: Auto-correlação de retornos do S&P500 em um gráfico semi-log onde é possível observar o comportamento exponencial ($e^{-\tau}$) com coeficiente de decaimento da ordem de $\tau = 4$ min. Adaptada de [54]



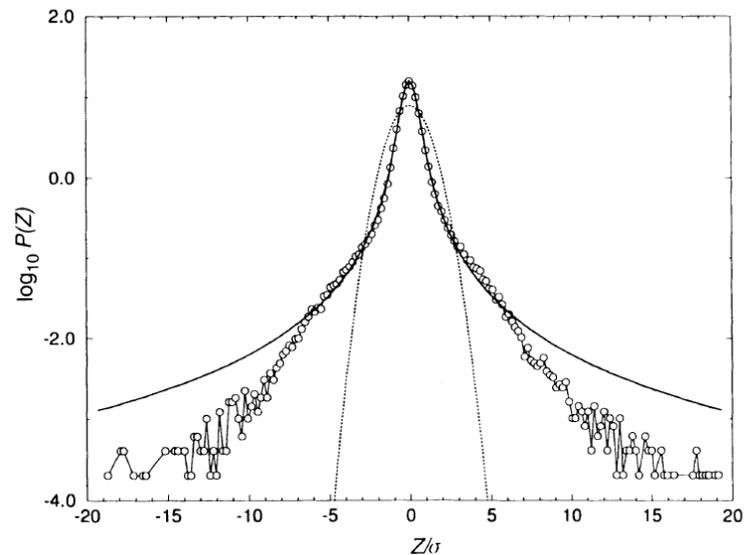
Fonte: Mantegna, R e Stanley, H, 2000 [54]

Figura 4.2: Auto-correlação de volatilidade do S&P 500, decaindo com uma lei de potência do tipo $t^{-0,3}$. Adaptada de [54]



Fonte: Mantegna, R e Stanley, H, 2000 [54]

Figura 4.3: Distribuição de probabilidade para os retornos definidos como Z do S&P500, onde pode-se notar a presença de caudas grossas, ou seja, probabilidades não nulas de eventos extremos e a curva gaussiana (em pontos) não reflete o comportamento dos dados. Adaptada de [54]

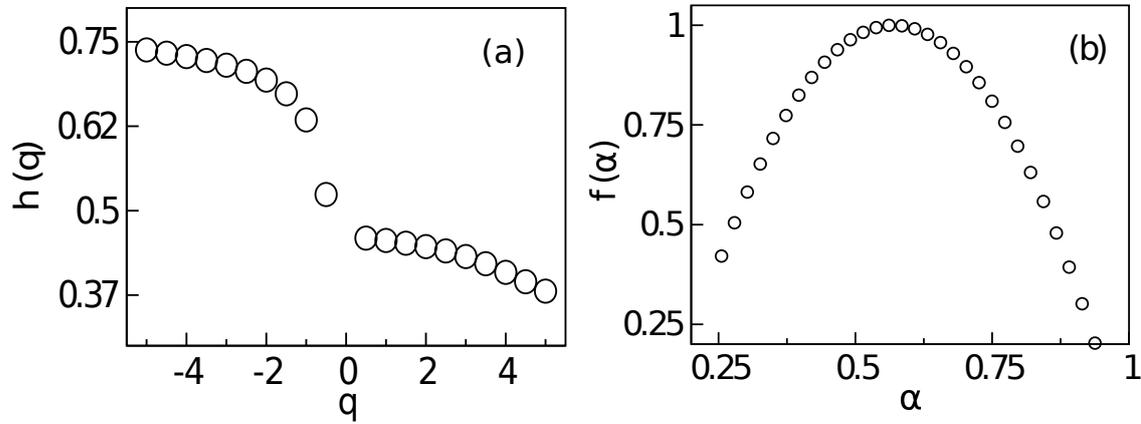


Fonte: Mantegna, R e Stanley, H, 2000 [54]

Por essas tantas propriedades, as séries temporais oriundas de retornos em sistemas econômicos, são ótimos candidatos ao estudo da aplicação dos teoremas gerais de convergência à gaussiana, TLC e TSL. Por não constituírem em um sistema i.i.d., eles não satisfazem a condição do TLC, mas o TSL sugere que mesmo assim poderiam em princípio evoluírem para uma distribuição normal, com a agregação de desvio padrão.

A análise de convergência dessas séries, também mostra um carácter de observação ímpar acerca das próprias propriedades inerentes à esses sistemas, de tal forma que os parâmetros determinados durante a aplicação dos teoremas, poderá elucidar algumas

Figura 4.4: (a) Dependência do expoente de Hurst com o q -ésimo momento. (b) Espectro multifractal em função do expoente de α , que no caso é conhecido como expoente de Hölder.[40]



Fonte: Nascimento, C.M. e colaboradores, 2008, [40]

questões ainda pertinentes, como por exemplo o papel da multifractalidade no contexto de correlações de longo alcance, ou mesmo a própria dinâmica estocástica de mercado.

5

O TLC e o TSL aplicadas a séries financeiras

Existem muitos problemas em aberto no âmbito da dinâmica de sistemas complexos, e isso tem sido o foco de diversas pesquisas [58, 59]. Flutuação em sistemas como esses normalmente não seguem uma distribuição gaussiana, são carregados de correlações de curto e longo alcances, como por exemplo sistemas financeiros [47, 49, 53, 54, 60, 61], como discutido no capítulo anterior.

O objetivo desse capítulo é aplicar as ferramentas de análise e verificar como o TSL se comporta ao ser aplicado em sistemas correlacionados como os sistemas econômicos.

Como é conhecido, o teorema do limite central afirma que a distribuição da soma de N variáveis aleatórias sem correlação ou fracamente correlacionadas converge para uma distribuição Gaussiana a medida que N cresce e tende ao infinito. Já o Teorema das Seções de Lévy garante essa mesma convergência para distribuições gaussianas mesmo em regime fortemente correlacionado, com base no parâmetro t da seção. Qual será o preço pago para que a convergência via TSL aconteça? Nesse capítulo serão mostrados alguns resultados que apresentam uma multifractalidade residual presente nas séries mesmo depois de alcançada a convergência onde é suposto que o sistema está no regime gaussiano quando os momentos estatísticos, no caso o terceiro e o quarto, atingem valores canônicos para a distribuição gaussiana, como era de se prever ao uso de tais teoremas.

Para garantir a robustez do trabalho, foram aplicados vários testes para determinar o ponto de convergência para o sistema. Os testes foram, o cálculo dos momentos estatísticos, a multifractalidade do sistema, através do espectro multifractal e do expoente de Hürst e por fim os testes estatísticos de comparação, o teste KS e o teste KL de divergência.

Alguns dos resultados aqui expostos, mais precisamente as análises de convergência dos momentos estatísticos e a análise multifractal presentes na seção 5.2, foram publicados na revista EPL[16].

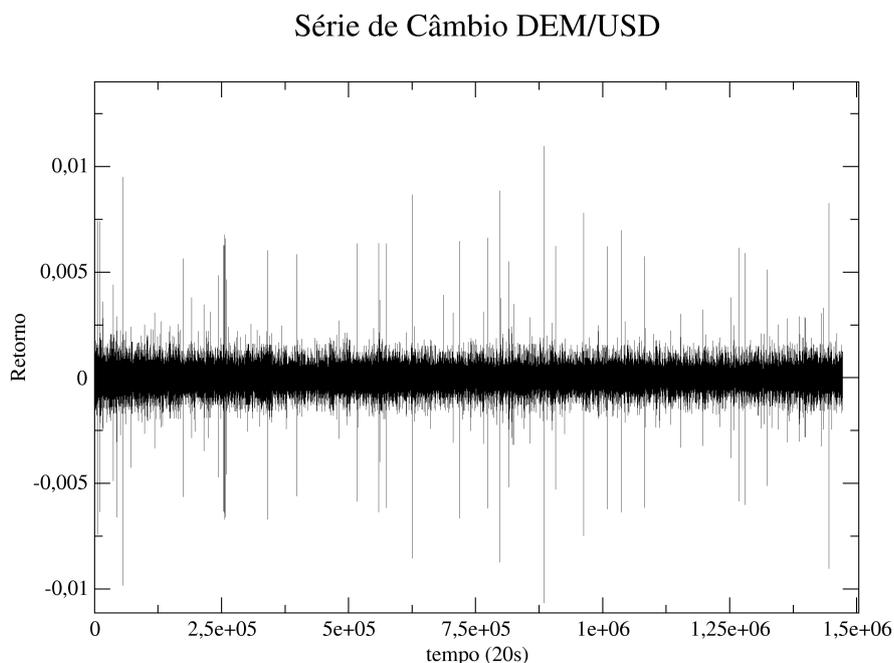
5.1 O banco de dados

As análises são feitas em um banco de dados de uma série temporal de câmbio de DEM/USD(Deutsch Mark/ US Dollar), ou seja, o Marco Alemão frente ao Dólar americano. Essas séries são do tipo *tick-by-tick*, um termo econômico que significa que são dados de alta frequência, ou seja, fornecidos em intervalos de tempo bem pequenos, elas foram fornecidas por Olsen & Associates e compreendem um período de um ano,

começando em primeiro de Outubro de 1992, até 30 de Setembro de 1993. A série tem 1472240 pontos o que significa que é em média 1 ponto a cada vinte segundos de transações.

A figura 5.1 mostra como a série de retornos se comporta. É uma série clássica de estudos de séries financeiras, e como tal possui todas as propriedades descritas no capítulo anterior.

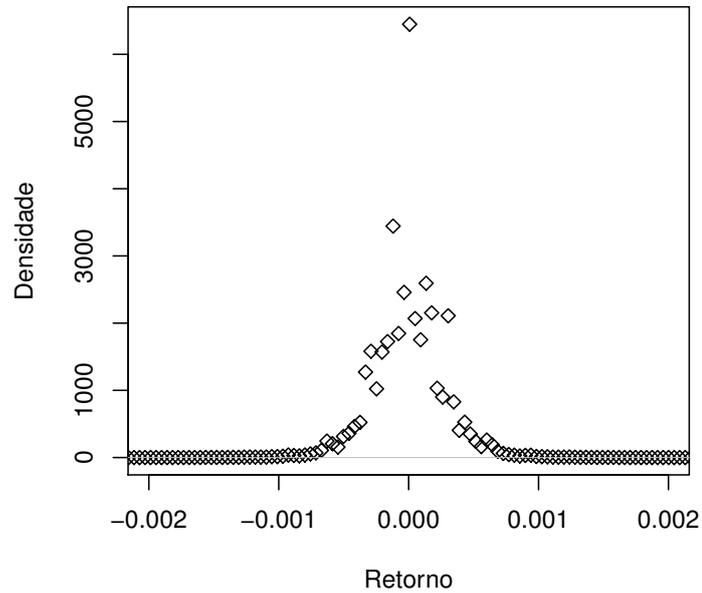
Figura 5.1: Série de retornos do câmbio do DEM/USD fornecida por Olsen & Associates



Fonte: Olsen & Associates, 1993

A distribuição dos retornos dos dados da série é de extrema importância. Na figura 5.2 observa-se um gráfico dessa densidade que claramente não é uma distribuição gaussiana. O fato da serie dispor de muitos dados faz com que a distribuição fique muito leptocúrtica, ou seja, com a *kurtosis* bem grande, fornecendo uma distribuição “afinada” no centro. Para melhor observar algumas propriedades, tais como as caudas grossas, possível analisar o comportamento dos dados utilizando apenas 10^4 pontos da série. Essa distribuição pode ser útil para observar a sutileza das flutuações.

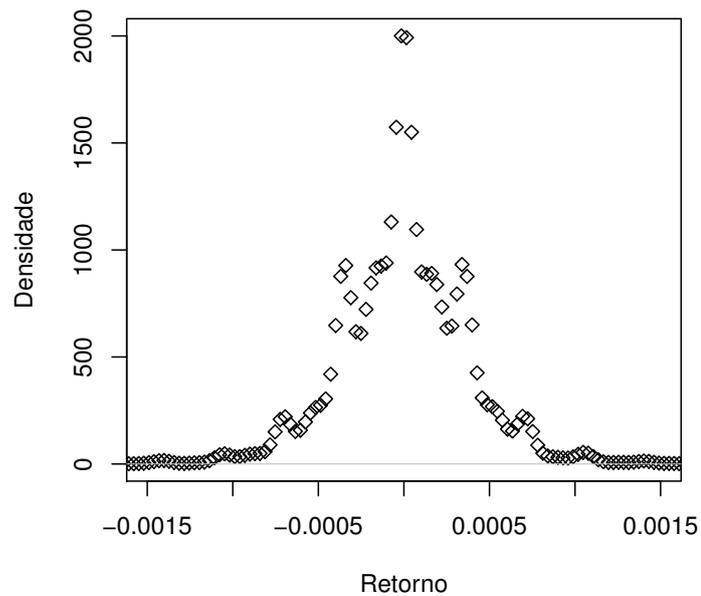
Figura 5.2: Densidade dos retornos da série
Distribuição dos dados



Fonte: Autor, 2014.

Outro ponto importante das séries temporais econômicas e extremamente relevantes para o estudo de convergência, são as correlações na série, tanto a correlação dos retornos, quando a correlação na volatilidade. Essa última é que fornece as correlações de longo al-

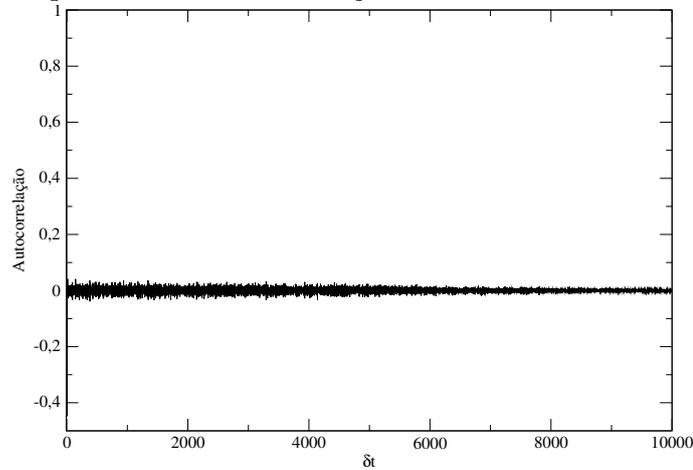
Figura 5.3: Densidade dos retornos da série reduzida para 10^4 dados
Distribuição dos dados



Fonte: Autor, 2014.

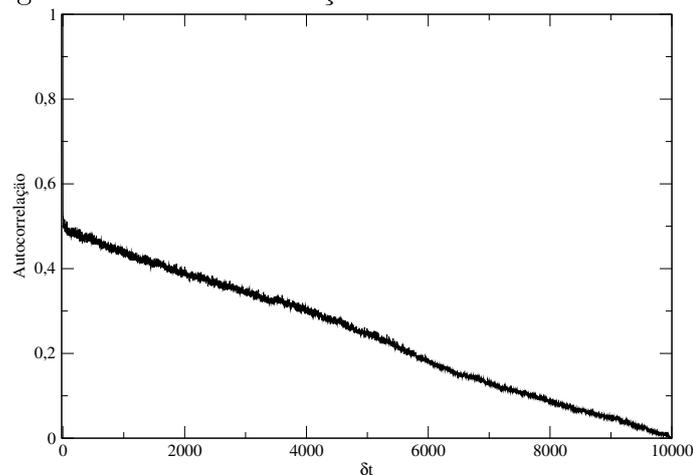
cance da série, e faz com que o sistema não entre em convergência pelo TLC simplesmente, apesar da série em si, ser descorrelacionada, como pode ser visto na figura 5.4.

Figura 5.4: Auto-correlação das séries de retorno



Fonte: Autor, 2014.

Figura 5.5: Auto-correlação das séries de volatilidade

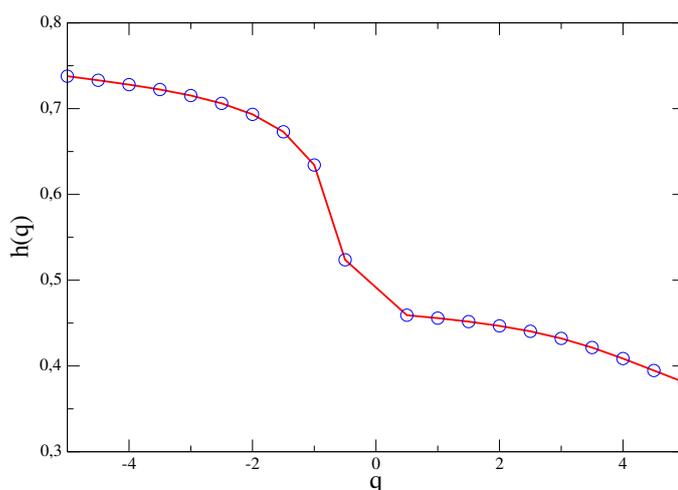


Fonte: Autor, 2014.

Nas figuras 5.4 e 5.5 observa-se os chamados fatos estilizados, as séries de retorno são completamente descorrelacionadas enquanto as séries de volatilidade possuem correlações de longo alcance.

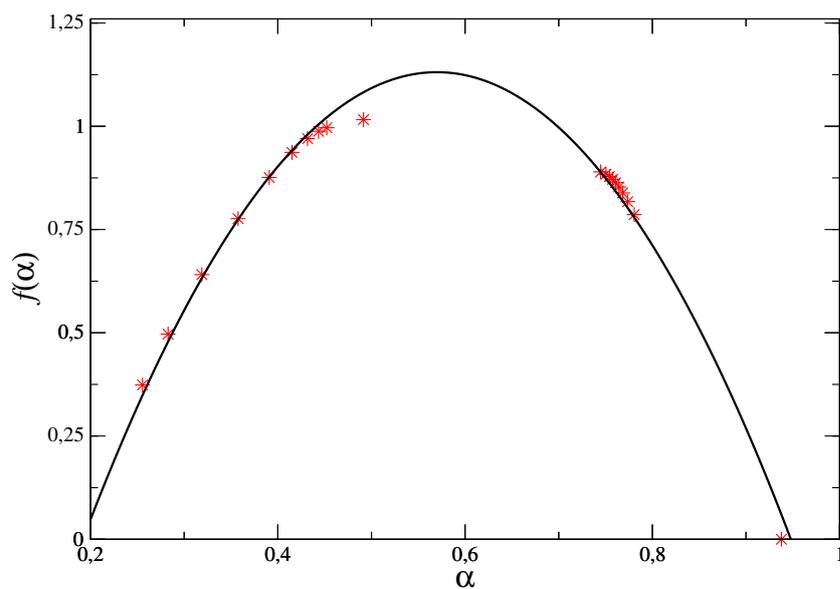
Outro fato estilizado importante é a multifractalidade da série. Nas figuras 5.6 e 5.7 esse aspecto se faz presente sendo um dos principais fatores de discussão mais a frente. O expoente de Hürst é extremamente dependente do momento calculado e o espectro multifractal mostra-se com uma curva bastante larga, o que acentua a evidência de propriedades multifractais do sistema.

Figura 5.6: Dependência do expoente de Hürst com o q -ésimo momento pelo método MF-DFA



Fonte: Autor, 2014.

Figura 5.7: Comportamento do espectro multifractal, calculado com as equações 2.138 e 2.139



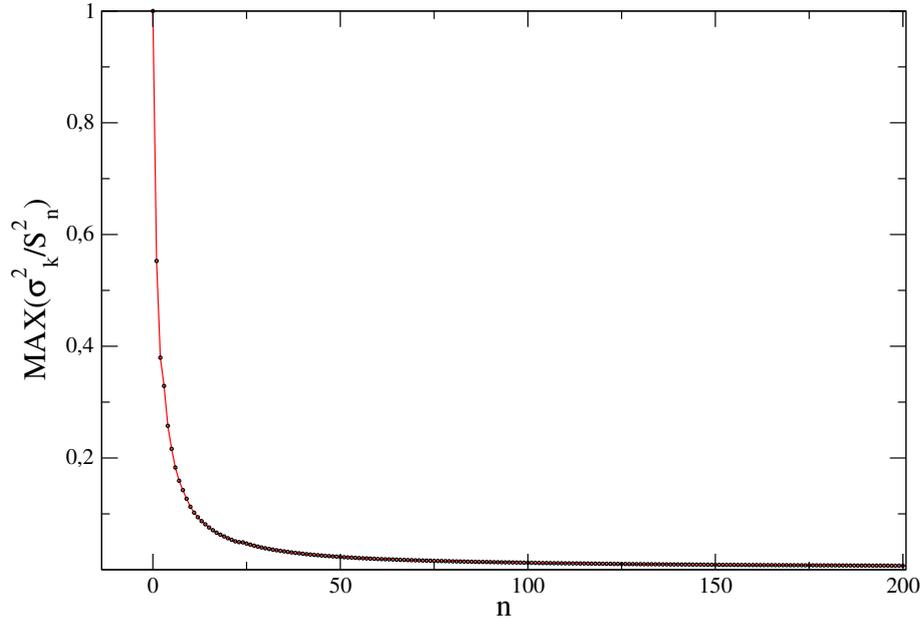
Fonte: Autor, 2014.

5.2 A convergência

Antes de verificar a aplicação dos teoremas, afim de verificar a validade do TSL para esse dado sistema, se faz necessário a verificação da obediência à condição de Lindeberg para a série. A figura 5.8 mostra o comportamento da série temporal em face da condição de Lindeberg.

Para a análise da convergência, dado o fato de que a série respeite a condição de Lindeberg, foram geradas diversas séries com base no processo pelo qual os teoremas devem ser executados e com cada uma das novas séries determinadas, serão calculadas

Figura 5.8: Condição de Lindeberg para a série empírica do DEM/USD, mostrando que a série tem uma variância finita e por esse motivo os teoremas podem ser aplicados à ela.



Fonte: Autor, 2014.

suas propriedades e comparadas com a gaussiana, para medir o grau de convergência da série.

Para o TLC basta aplicar o conceito mostrado na equação 3.1, que expressa como devem ser agregadas as variáveis.

No caso do TSL é preciso ter um pouco mais de cuidado, pois o que é agregado no TSL são as variâncias locais e tudo depende o parâmetro t da seção. A análise começou com uma seção inicial de $t = 10^{-13}$ com pequenos incrementos de $\Delta t = 2,5 \cdot 10^{-7}$, que são suficientemente ínfimos para que exista uma variação arbitrariamente suave nos dados.

Afim de comparar as convergências ao regime gaussiano fornecido por cada teorema é definido uma variável de tempo chamada de τ onde é possível analisar a “velocidade” com que se atinge a convergência da série. Tal parâmetro é de difícil análise tendo em vista a diferente natureza que envolve a convergência via TLC e TSL. Para o TLC foi utilizado como tempo de convergência, o número de variáveis agregadas. No caso do TSL o tempo τ é definido como

$$\tau_t = \frac{\sigma_{S_t}^2}{\sigma_{S_n}^2}, \quad (5.1)$$

que é a razão entre a variância da seção t e a variância da série original.

Como primeira análise é importante observar como os momentos, *kurtosis* e *skewness*, se comportam tanto com a agregação via TLC e com a aplicação do TSL. Como comparativo foi gerada uma série de ruído branco, uma série de variáveis aleatórias que são decorrelacionadas e com variáveis i.i.d e por tanto satisfaz a condição de Lindenberg e ao TLC convergindo rapidamente para uma distribuição gaussiana e sendo um excelente parâmetro para compararmos as séries do DEM/USD. O ruído branco também foi sujeito aos dois teoremas para mais uma vez constatar que ambos são capazes de nos fornecer os valores gaussianos, dada uma série i.i.d..

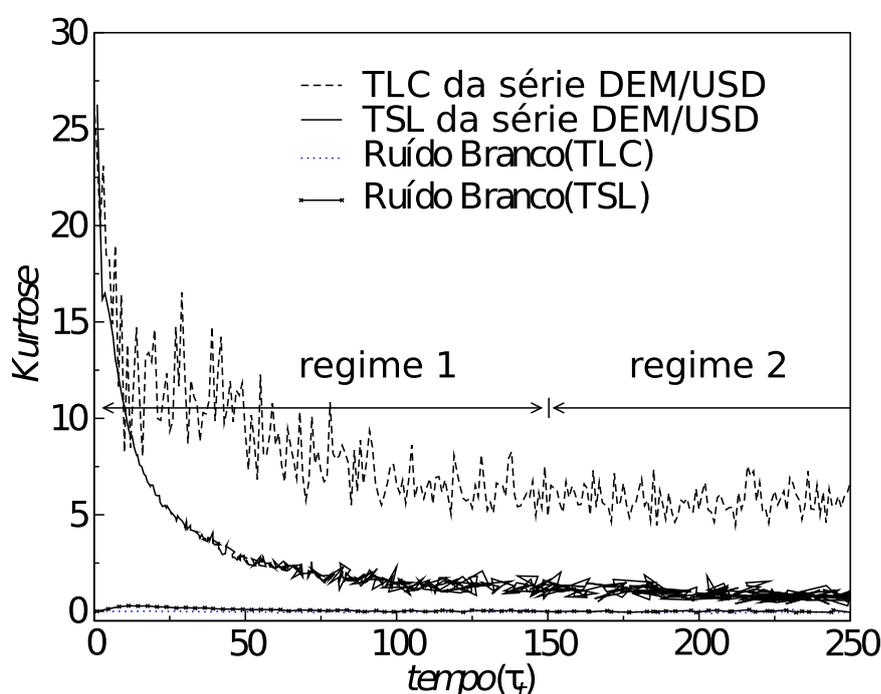
Para garantir a robustez e precisão da análise os valores iniciais para a *kurtosis* e *skewness* para a série original do DEM/USD são:

Tabela 5.1: Valores da *skewness* e *kurtosis* para a série original do DEM/USD

	Valores iniciais
<i>skewness</i>	$2,6799 \cdot 10^{-02}$
<i>kurtosis</i>	26,3767

Fonte: Autor, 2014.

Figura 5.9: As curvas de convergência da *kurtosis* para os dados analisados



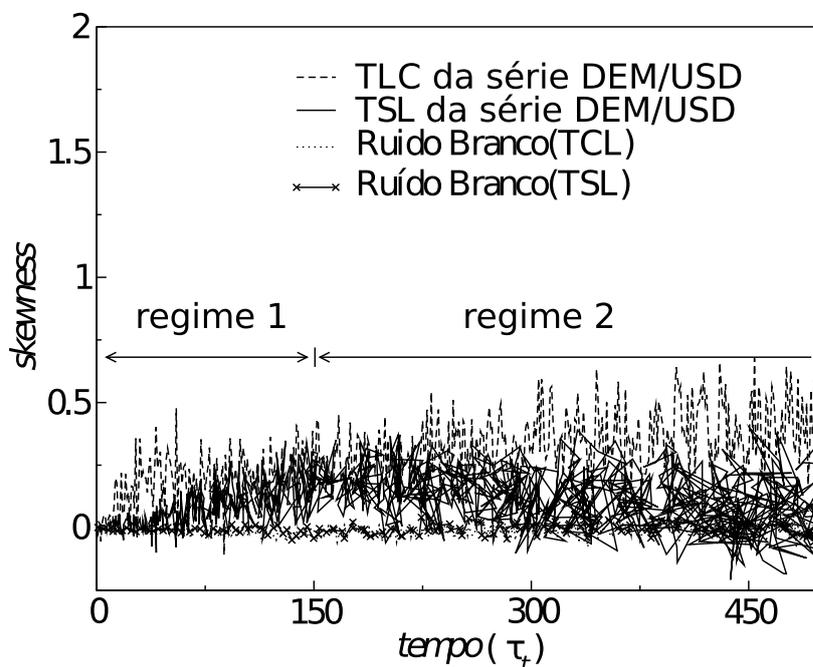
Fonte: Autor, 2014.

Na figura 5.9 é destacado que existem dois regimes presentes. O primeiro regime trata da convergência propriamente dita e se inicia em $\tau = 0$ e segue até $\tau = 150$ onde é atingido um regime que estatisticamente se mostra estacionário para a convergência da *kurtosis* em ambos os teoremas. É fácil observar que o TSL apresenta uma convergência muito mais enfática e rápida em comparação ao TLC. Apesar disso, é importante frisar que essa convergência se dá por meio de um tempo próprio definido pela equação 5.1 e não em número de variáveis agregadas. Como discutido em capítulos anteriores o TSL agrega as variáveis dependendo da variância local delas, podendo agregar muitas ou poucas variáveis dependendo apenas do parâmetro t da seção.

A *kurtosis* é um bom “termômetro” para medir a convergência, isso porque um dos fatos estilizados dos mercados financeiros, discutidos no capítulo anterior, é que a distribuição para as séries de retornos são leptocúrticas, ou seja, com uma *kurtosis* maior do que a

gaussiana. A tendência da *kurtosis* seguir para zero indica que a distribuição da nova série se torna mais arredondada.

Figura 5.10: As curvas de convergência da *skewness* para os dados analisados



Fonte: Autor, 2014.

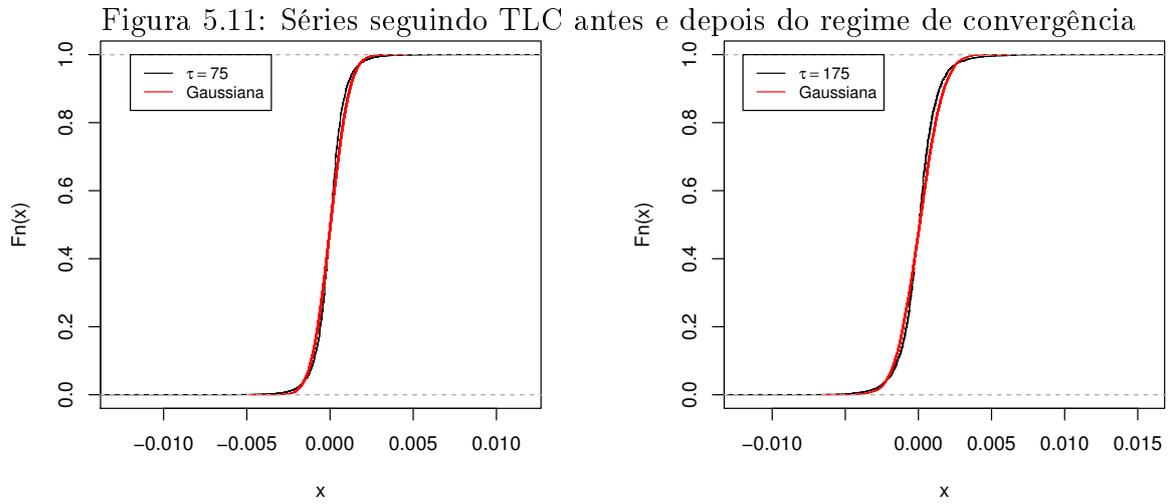
A curva de *skewness*, apresentada na figura 5.10, é muito mais ruidosa mas bem próxima do valor da *skewness* da gaussiana, ou seja a nulidade, isso porque a gaussiana é uma distribuição simétrica e a *skewness* mede um dos graus de assimetria. Tendo em vista que a distribuição apresentada nas figuras 5.2 e 5.3, que é a distribuição original da série, também ser uma distribuição visualmente simétrica e isso foi demonstrado na tabela 5.1, era de se esperar a rápida convergência.

Além das análises dos momentos, foram utilizados alguns testes estatísticos que trazem muita confiança com relação a convergência da série, os testes KL de divergência e KS. Tais testes serão úteis para determinarmos tanto se a escolha do ponto de convergência está correto, quando se a série realmente convergiu e se aproximou de uma distribuição gaussiana, como afirmam os dois teoremas aqui confrontados.

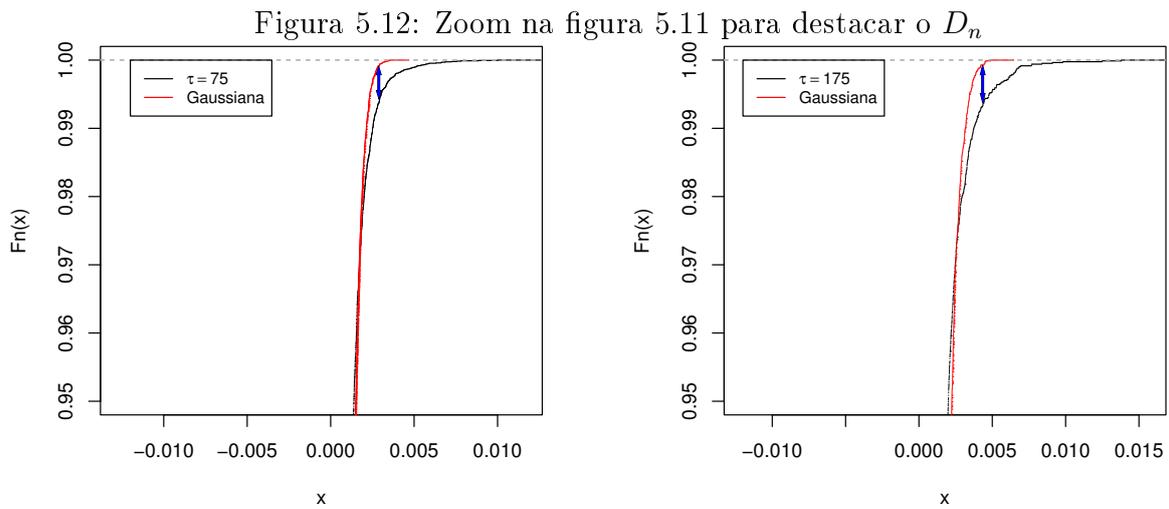
O primeiro teste a ser utilizado é o teste KS que consiste em um teste não-paramétrico onde um parâmetro D_n é calculado conforme mostrado na equação 3.64. Na presente análise foi utilizado o programa de estudos estatísticos R[20] e com ele e algumas sub-rotinas presentes no programa, foram geradas séries com dados distribuídos gaussianamente e comparados um a um com as séries a medida que o tempo τ , o tempo já definido, crescia.

Nas figuras 5.10,5.11,5.12 e 5.13 é observado de forma pictórica como as análises se comportavam ao comparar FDAE gerada tanto por TLC quanto por TSL com as séries geradas pelo programa R distribuídas gaussianamente, com a mesma média das séries empíricas e o mesmo desvio padrão, pois, o teste consiste em comparar duas FDAE de séries e por isso além da série do DEM/USD é gerada artificialmente uma série que tem

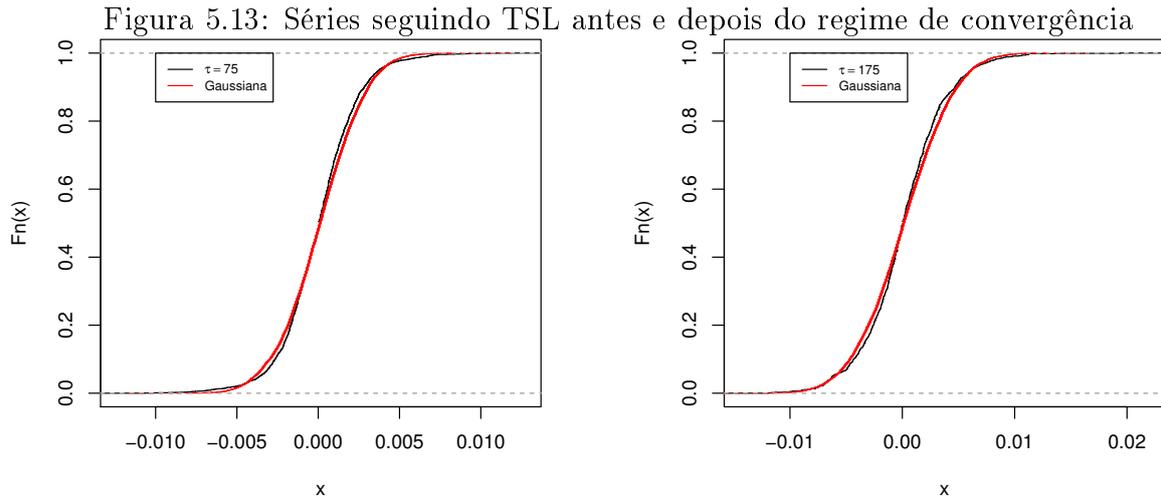
como distribuição uma curva normal. Para esse comparativo é tomado para cada teorema duas séries que estão em regimes diferentes, propostos pela convergência da *kurtosis*.



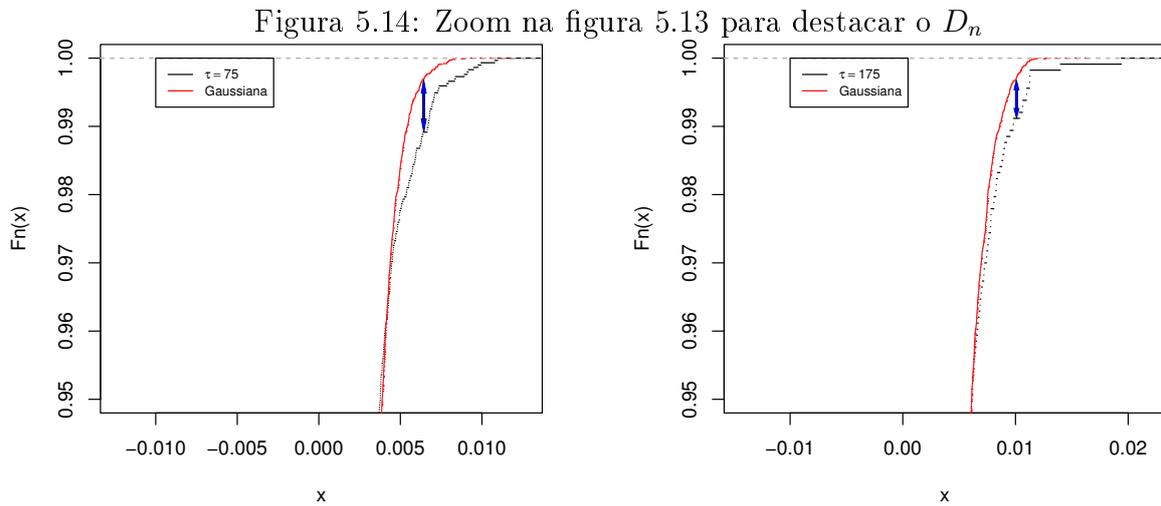
Fonte: Autor, 2014.



Fonte: Autor, 2014.



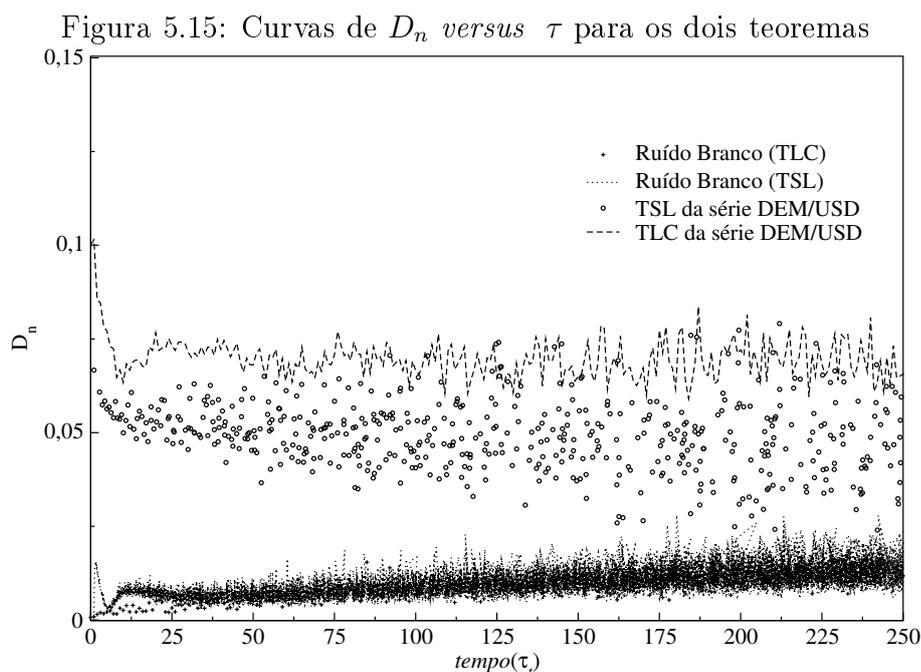
Fonte: Autor, 2014.



Fonte: Autor, 2014.

É possível observar nessas figuras dois fatos importantes, que as distribuições tem D_n muito pequenos e mesmo assim os valores para o TLC parecem menor do que os valores para o TSL. Um outro ponto que deve ser verificado é a diminuição do D_n a medida que o τ cresce. Visualmente é possível enxergar isso, pois as escalas são as mesmas nos dois casos.

Para confirmar essas suspeitas de comportamento, foi feito um gráfico de D_n por τ para todos os sistemas, afim de poder compreender o mecanismo da convergência das séries, via o teste KS.



Fonte: Autor, 2014.

A figura 5.15 torna visível a convergência da série. Apesar de um pouco ruidosa as curvas mostram que um padrão de convergência é rapidamente alcançado entretanto não se mantém constante. Sendo assim algo surge em questionamento, através das curvas de D_n não é possível determinar uma clara transição de comportamento e sim, uma convergência para um valor médio estacionário com o avanço do tempo. Na tabela 5.2 é possível conferir os valores médios para a aplicação de cada um dos teoremas e confirmar que a convergência não é obtida, pois o parâmetro D_n é não nulo, o que indica que existe alguma diferença entre as FDAEs da série do DEM/USD aplicado sobre os teoremas e um regime puramente gaussiano.

Tabela 5.2: Tabela dos valores médios de D_n

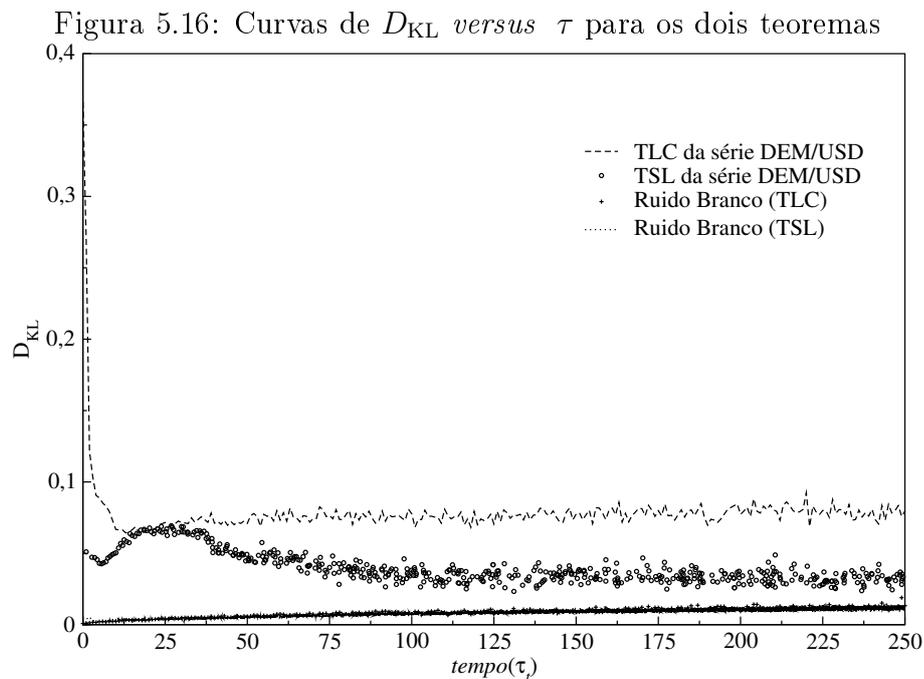
Teorema	D_n
TLC	0,073
TSL	0,052

Fonte: Autor, 2014.

Apesar de apresentar um regime tendencioso, o teste KS fornece uma nova argumentação, pois um gráfico bastante ruidoso como o mostrado leva a crer que possivelmente a convergência não foi expressamente atingida. Algo impede que tal fato se concretize, possivelmente correlações que ainda persistem em existir no sistema, mesmo após a aplicação dos teoremas.

Outro teste importante e discutido em capítulos anteriores é o teste KL de divergência. Esse teste mede a entropia de Shannon cruzada entre as duas distribuições, e a medida da informação que é perdida ao transformar uma distribuição na outra, o que será um excelente medidor de convergência, pois um parâmetro nulo para esse teste indica que nenhuma informação é perdida da conversão das distribuições, o que implica que os dados são provenientes da mesma distribuição.

Para o cálculo do método KL também foi utilizado o *software* R de cálculos estatísticos[20]. Como no caso do teste KL, foi gerada uma série de números distribuídos gaussianamente para serem usados como base comparativo para o teste. Diferente do teste KS que mede o parâmetro D_n usando a FDAE, o teste KL utiliza a própria distribuição de densidade de probabilidade para realizar o teste comparativo, distribuição essa facilmente calculada utilizando o programa R.



Fonte: Autor, 2014.

Tabela 5.3: Tabela dos valores médios de D_{KL} em cada um dos teoremas, mostrando que mesmo sendo da ordem de 10^{-2} ainda não são nulos

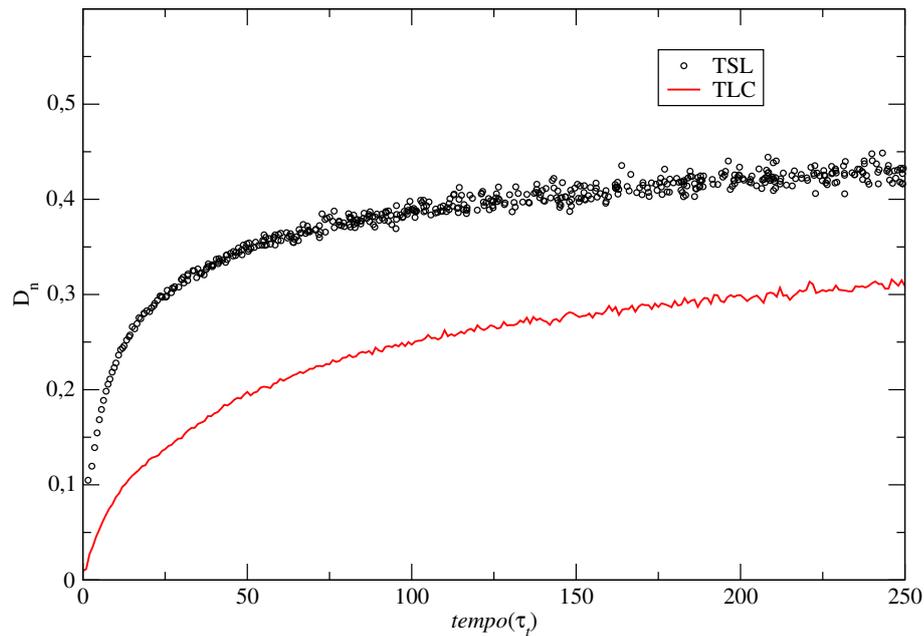
Teorema	D_{KL}
TLC	0,0857
TSL	0,0380

Fonte: Autor, 2014.

Na figura 5.16 e a tabela 5.3, observam-se curvas muito menos ruidosas do que o apresentado na figura 5.15. No caso do teste KL é possível inferir a pouca perda de informação entre uma curva gaussiana normal e uma produzida pelos dois teoremas, ou seja, um processo de convergência bem claro e com um parâmetro D_{KL} muito pequeno e próximo de zero, mesmo para valores pequenos de τ .

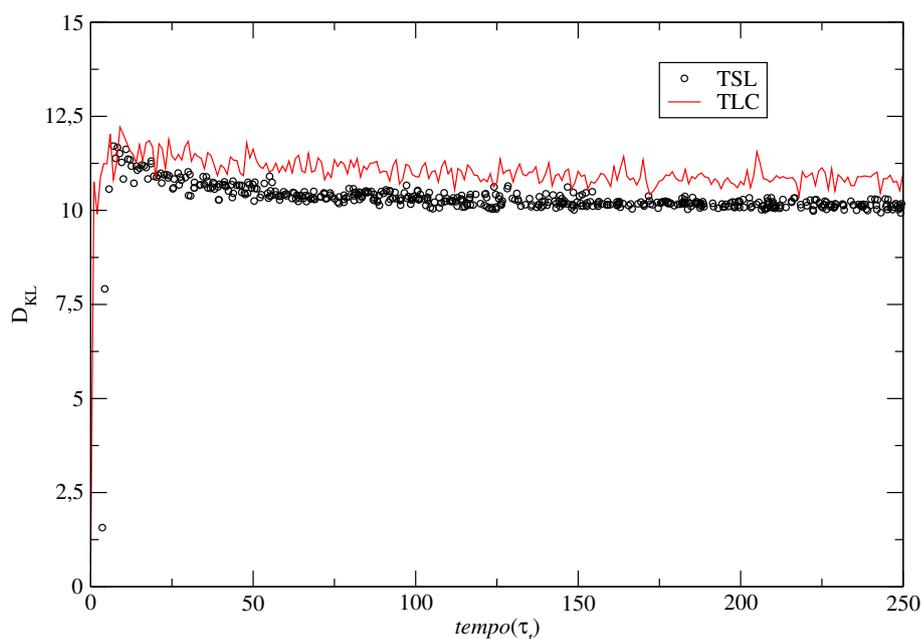
Utilizando os testes aqui apresentados, além de analisar a convergência para a série gaussiana, pode-se também verificar o quão distante as séries estão da série original do DEM/USD, sendo isso uma outra medida tanto do grau de correlação perdido quanto da divergência geral aplicada pela ação dos teoremas.

Figura 5.17: Teste KS comparando a série original do DEM/USD com a aplicação dos Teoremas.



Fonte: Autor, 2014.

Figura 5.18: Teste KL comparando a série original do DEM/USD com a aplicação dos Teoremas.



Fonte: Autor, 2014.

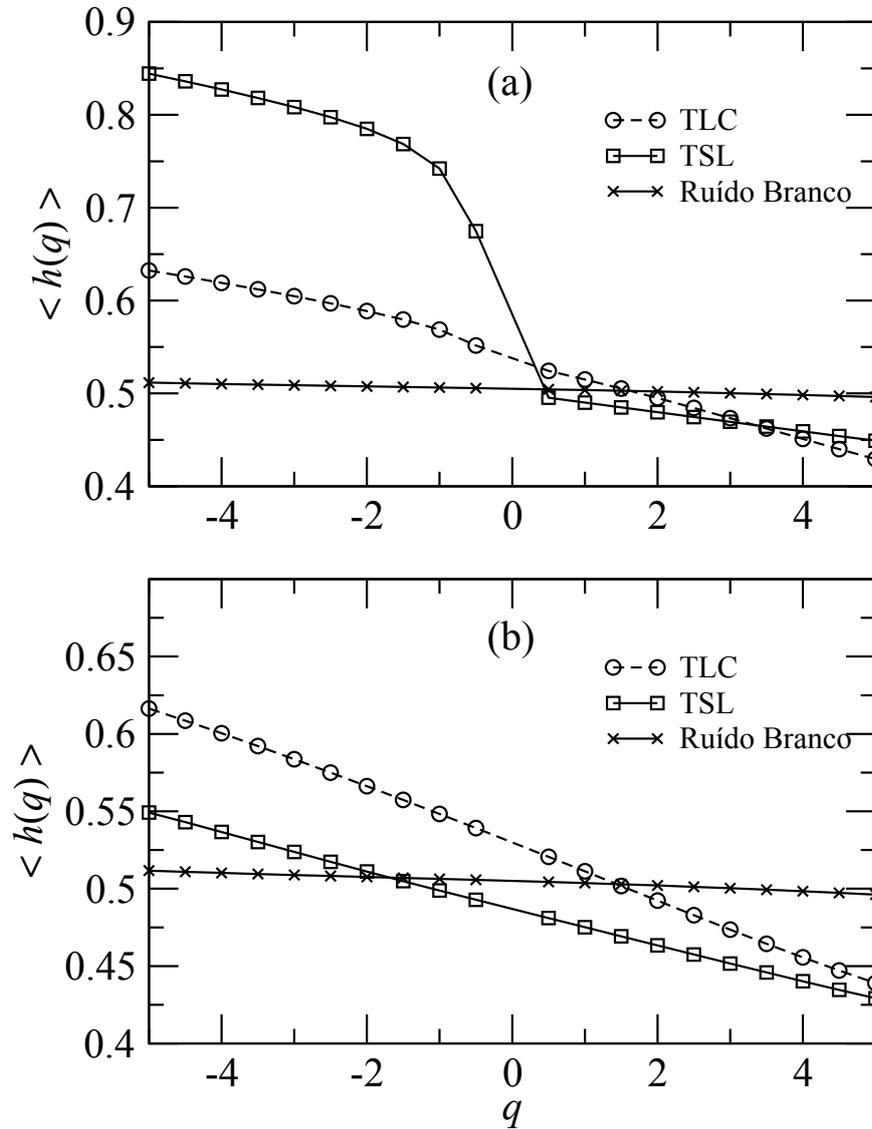
O distanciamento se torna evidente ao comparar as figuras 5.17 e 5.18, mostrando que a aplicação de ambos os teoremas são eficazes na quebra de correlações e transformando as séries antigas em novas séries que não representam todas as propriedades da série original do DEM/USD. Na figura 5.17 existe uma tendência clara de crescimento monotônico de D_n com o TSL se mostrando superior ao se distanciar da série original do DEM/USD. Enquanto isso, na figura 5.18 o comportamento observado é de quase de estacionariedade imediata do parâmetro D_{KL} nos dois casos, no TLC e para TSL onde ambos tem praticamente o mesmo valor.

Durante as análises é fácil perceber a clara convergência a um regime estacionário para as duas séries, tanto a série TLC quando a série TSL, ou seja, após algumas aplicações dos teoremas, as séries vão se alterando cada vez menos em aplicações consecutivas. Entretanto esse regime estacionário não condiz com os valores gaussianos para a *kurtosis* e para a *skewness* e nem mesmo atingindo o valor zero para os parâmetros D_n e D_{KL} dos testes KS e KL respectivamente. Algo então ainda reside nas entranhas da configuração das séries, alguma correlação residual ou algo mais do que isso, ou momentos de mais alta ordem que não tenham convergido efetivamente ao comportamento gaussiano.

Com o intuito de se aprofundar nas causas da falta de convergência das séries, é possível fazer uma análise do espectro multifractal, uma propriedade claramente observada na série original do DEM/USD (figuras 5.6 e 5.7) e que pode fornecer uma luz acerca dessa convergência para valores não gaussianos.

A figura 5.19 apresenta o comportamento do expoente de Hürst médio para o TLC e para o TSL em ambos os regimes, de convergência e estacionário. Para o primeiro caso o expoente de Hürst tem uma dependência completamente não-linear com o q-ésimo momento, mostrando um padrão multifractal muito claro. Quando existe a migração de regime a dependência não-linear do expoente passa a ser uma dependência linear mas

Figura 5.19: Expoente de Hürst médio em função do momento q . (a) Média para $\tau \leq 150$; (b) Média para $\tau > 150$



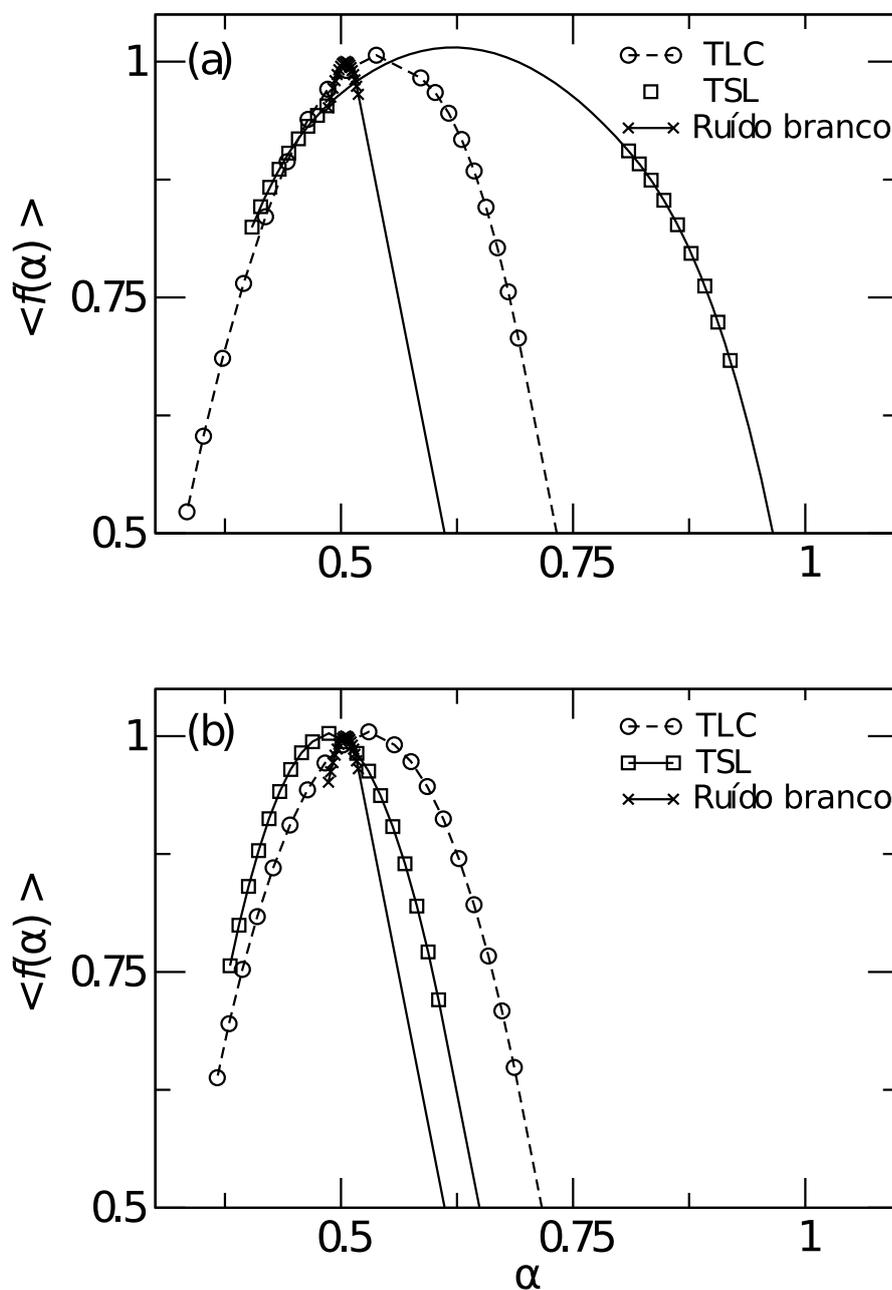
Fonte: Autor, 2014.

mesmo assim, $h(q)$ não assume um valor constante, e igual a $1/2$ como no caso gaussiano, e portanto ainda reside uma característica multifractal.

O mesmo padrão descrito para o expoente de Hürst pode ser muito bem observado na figura 5.20, onde apesar de ter uma redução de largura, quando ocorre a transição do regime de convergência para o regime estacionário, o sistema ainda mantém a característica multifractal com um espectro largo em comparação ao ruído branco, que aqui representa os valores da distribuição gaussiana.

A clara mudança de regime, observando o caráter multifractal, pode ser explicada com a quebra de correlações presentes na série. Esse fato é bem expresso nas figuras 5.17 e 5.18, onde foi medido usando o teste KL o quanto de informação é perdida em comparação a série original do DEM/USD. Mesmo com essa quebra de correlações o

Figura 5.20: Espectro multifractal, $f(\alpha)$, em função de α . (a) Média para $\tau \leq 150$; (b) Média para $\tau > 150$



Fonte: Autor, 2014.

regime estacionário não apresenta valores gaussianos em nenhum dos testes apresentados. Algumas explicações podem ser dadas para esse fator, a principal delas são correlações de longo alcance presentes nas séries financeiras. Outra explicação passível de ser aceita é sobre o fato de estar se tratando de um conjunto finito de dados e ambos os teoremas são aplicados para um conjunto infinito de dados, ou seja, a finitude dos termos da série temporal impede que a convergência ao regime gaussiano seja alcançada.

6

Traço do movimento Browniano fracionário

O movimento browniano foi descrito em 1827 pelo botânico Robert Brown enquanto observava em um microscópio grãos de pólen de *Clarkia pulchella* suspensos em água [62]. O movimento errático realizado por essas partículas suspensas foi descrito por Einstein em 1905, considerado seu *Annus mirabilis*, utilizando a teoria cinética de fluidos e a controversa, na época, mecânica estatística de Boltzmann[17]. Mandelbrot, como discutido no capítulo 4, utilizou o formalismo apresentado pelo movimento browniano e o aplicou em séries econômicas.

Mandelbrot assim, formulou uma generalização do conceito de movimento browniano, chamado movimento browniano fracionário[26, 63]. Além das aplicações em economia, Mandelbrot e Van Ness, co-autor do trabalho em destaque, observaram que sua generalização também, poderia ser aplicada em estudo de flutuações em sólidos, ou ruídos em sólidos. Algumas dessas flutuações são chamadas de “ruídos $1/f$ ” pelo fato da sua densidade espectral se comportar tomando uma forma do tipo ω^{1-2H} , onde ω é a frequência de oscilação e H é o chamado expoente de Hürst.

6.1 Definição do movimento browniano fracionário

A definição apresentada por Mandelbrot e Van Ness trata de algumas propriedades que as variáveis aleatórias possuem. Seja t um tempo qualquer em que houve uma realização de uma variável aleatória x , onde x pertence a um espaço amostral X sobre o corpo dos reais.

O movimento browniano usual, $B(t, x)$ é uma função real aleatória com incrementos gaussianos independentes. Dessa forma a diferença $B(t_2, x) - B(t_1, x)$ tem média zero e variância dada por $|t_2 - t_1|$ de tal maneira que se os intervalos (t_1, t_2) e (t_3, t_4) não se sobreponham, $B(t_2, x) - B(t_1, x)$ e $B(t_4, x) - B(t_3, x)$, são independentes.

Sendo assim, definindo duas quantidades H e b_0 , de tal forma que $0 < H < 1$ e b_0 é um número real arbitrário, a função aleatória $B_H(t, x)$, é o movimento browniano fracionário reduzido, com parâmetro H e iniciando em b_0 para $t = 0$. Para $t > 0$, $B_H(t, x)$ é definida como:

$$B_H(t, x) - B_H(0, x) = \frac{1}{\Gamma(H + 1/2)} \left\{ \int_{-\infty}^0 [(t - s)^{H-1/2} - (-s)^{H-1/2}] dB(s, x) + \int_0^t (t - s)^{H-1/2} dB(s, x) \right\} \quad (6.1)$$

ou em uma forma mais simétrica:

$$B_H(t_2, x) - B_H(t_1, x) = \frac{1}{\Gamma(H + 1/2)} \left\{ \int_{-\infty}^{t_2} [(t_2 - s)^{H-1/2}] dB(s, x) - \int_{-\infty}^{t_1} (t_1 - s)^{H-1/2} dB(s, x) \right\} \quad (6.2)$$

Paul Lévy[64] já havia discutido brevemente sobre uma variável similar à $B_H(t, x)$, chamada Holmgren-Riemann-Liouville integral fracionada, dada por:

$$B_H^0(t, x) = \frac{1}{\Gamma(H + 1/2)} \int_{-\infty}^t [(t - s)^{H-1/2}] dB(s, x), \quad (6.3)$$

essa integral, valoriza em excesso o local da origem.

Entre as diversas propriedades que essas funções apresentam, em destaque está o fato de que o desvio padrão segue uma lei de potência, na forma:

$$\langle [B_H(t + T, x) - B_H(t, x)]^2 \rangle = 2DT^{2H}, \quad (6.4)$$

onde T é um incremento temporal, e D é a constante de difusão. Também é importante saber que essas funções são auto-similares, ou seja,

$$B_H(at, x) \sim |a|^H B_H(t, x), \quad (6.5)$$

e a correlação entre incrementos futuros e incrementos passados, com alguns parâmetros definidos com valor unitário através da escolha correta de unidades dimensionais, é dada por:

$$C(t) = \frac{\langle -B_H(-t, x)B_H(t, x) \rangle}{B_H(t, x)^2} = 2^{2H-1} - 1, \quad (6.6)$$

onde pode-se observar que o sistema é independente para $H = 1/2$, como é conhecido no movimento browniano usual.

Moura e Lyra[65] analisaram o comportamento do modelo de Anderson unidimensional[66] utilizando uma desordem correlacionada nas energias dos sítios da cadeia. Nesse trabalho eles definiram que o potencial deveria seguir uma densidade espectral a partir de um traço do movimento browniano fracionário, tomando a densidade como:

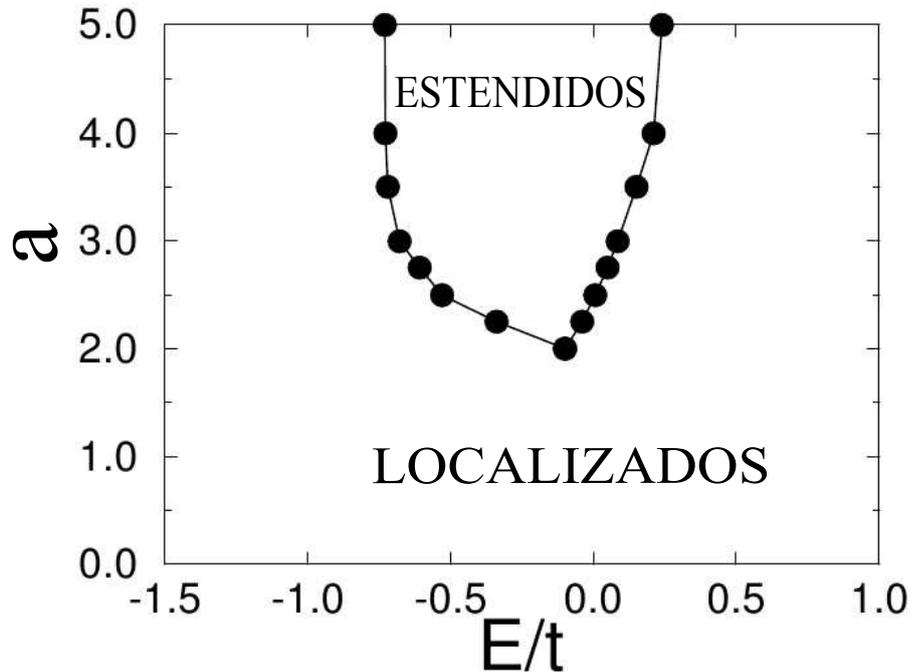
$$S(k) = \frac{1}{k^a}, \quad (6.7)$$

onde $S(k)$ é a transformada de Fourier da função de correlação entre dois sítios quaisquer na rede. O parâmetro a mede o grau de correlação da série, donde, para $a = 0$, a série é completamente descorrelacionada. Moura e Lyra perceberam também que para $a > 2$, o sistema pode sofrer uma transição metal-isolante, ao contrário de outros modelos anteriores unidimensionais, do tipo *dimmer*, que apresentavam apenas algumas faixas ressonantes de energia.

A série com as energias de cada um dos i sítios, de um total de N sítios, é gerada através da transformada de Fourier discreta[67, 68, 69, 70, 71] resultando na equação:

$$\epsilon_i = \sum_{k=1}^{N/2} [S(\omega_k)\Delta\omega_k]^{1/2} \cos(\omega_k t_i + \phi_k) \quad (6.8)$$

Figura 6.1: Diagrama de fase no plano $(E/t, a)$, mostrando os estados estendidos e localizados.



Fonte: De Moura, F.A.B.F. e Lyra, M.L., 1998, [65]

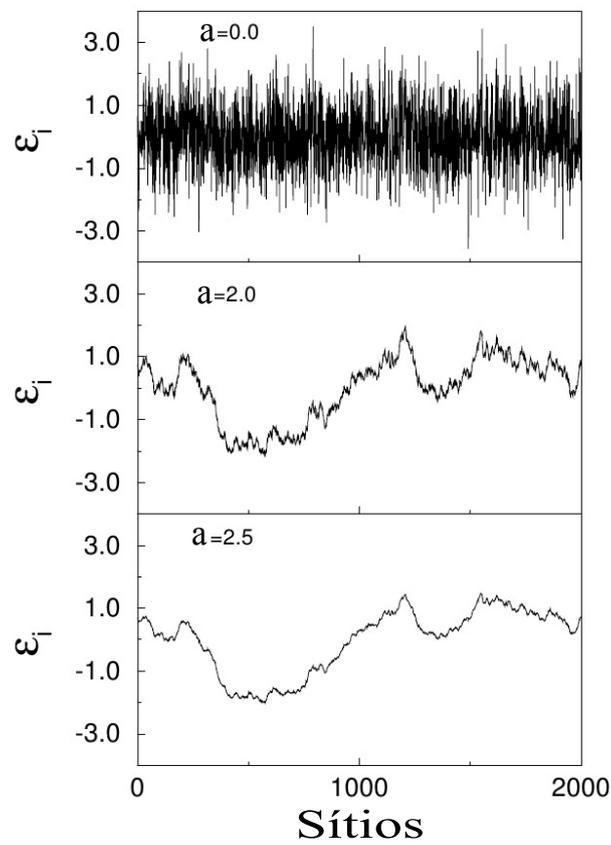
Considerando que as frequências ω_k são múltiplos da frequência fundamental $\Delta\omega = \frac{2\pi}{T}$ e que os ângulos de fase são distribuídos uniformemente sobre o intervalo $[0, 2\pi]$, obtêm-se para o espectro $S(\omega) = 1/\omega^a$ a série descrita como:

$$\epsilon_i = \sum_{k=1}^{N/2} \left[k^{-a} \left| \frac{2\pi}{N} \right|^{(1-a)} \right]^{1/2} \cos \left(\frac{2\pi i k}{N} + \phi_k \right), \quad (6.9)$$

que está normalizada de tal maneira que $\langle \epsilon \rangle = 0$ e $\sqrt{\langle \epsilon^2 \rangle - \langle \epsilon \rangle^2} = 1$. Pode-se observar o comportamento dessa série gerada pela equação 6.9 com valores diversos do parâmetro a na figura 6.2.

Por ser um resultado fisicamente interessante, sendo o primeiro a demonstrar a transição metal isolante em sistemas 1D, violando a lei de escala[66], o trabalho de Moura e Lyra já possui mais de 300 citações até os dias de hoje. Tendo como partida uma série correlacionada onde a correlação pode ser facilmente ajustada, esse modelo é suscetível à aplicação dos teoremas já aqui discutidos, o TLC e o TSL, afim de compreender qual o papel que as correlações implicam na convergência dessas séries e possivelmente entender melhor a transição apresentada por esse sistema.

Figura 6.2: Energias por sítios que foram geradas pela equação 6.9, com $N=2000$; onde $a = 0$ é uma sequência não correlacionada; $a = 2.0$ traço de um movimento browniano fracionário; $a = 2.5$ traço de um movimento browniano fracionário com incrementos de persistência.



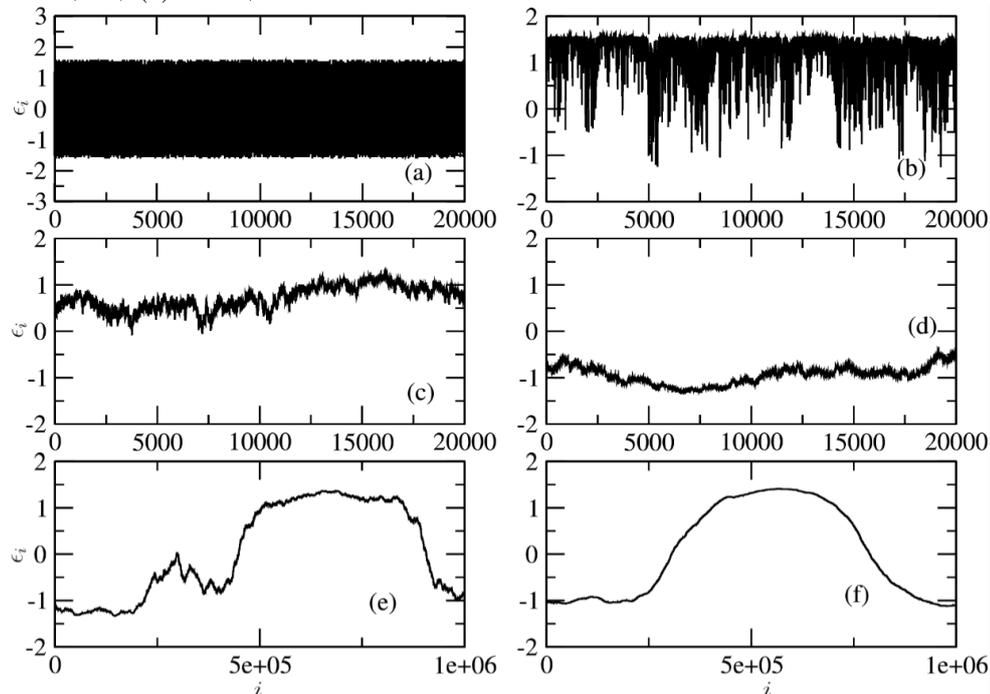
Fonte: De Moura, F.A.B.F. e Lyra, M.L., 1998, [65]

7

O TLC e o TSL aplicados à séries de sítios de energia correlacionados

Para as análises pertinentes acerca da importância e do papel que as correlações empregam na convergência das séries temporais, foram geradas algumas séries utilizando a correlação descrita no capítulo anterior. As séries foram geradas para diversos valores do parâmetro a , com 1048576 variáveis e com uma média nula, afim de possuir uma quantidade de dados próxima a que foi analisada no capítulo 5. Até a presente data, os resultados aqui abordados ainda são inéditos na literatura, não tendo nenhum outro trabalho publicado com a mesma abordagem ou conclusões.

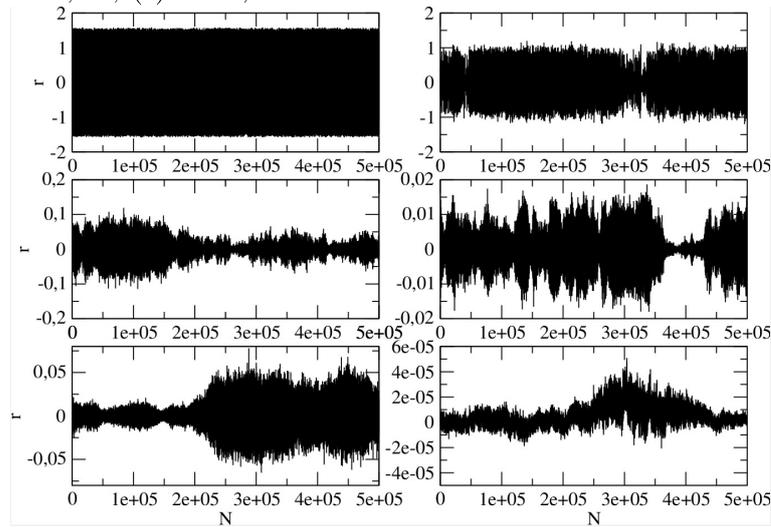
Figura 7.1: Várias séries geradas artificialmente com parâmetros diferenciados, com 2^{20} dados gerados pela equação 6.9. (a) $a = 0,00$, (b) $a = 1,00$, (c) $a = 1,50$, (d) $a = 2,00$, (e) $a = 2,50$, (f) $a = 3,00$



Fonte: Autor, 2014.

Essas séries não possuem padrões limitantes, o que impediria a princípio, ser capaz de observar algumas propriedades, tais como leis de escala e até mesmo a validade do teorema de Lévy. Sendo assim, afim de melhorar as análises, padronizar os resultados com base no capítulo 5, e tornar a série centrada na média, foi calculado o retorno dos dados, ou seja, o logaritmo da diferença de termos consecutivos da série.

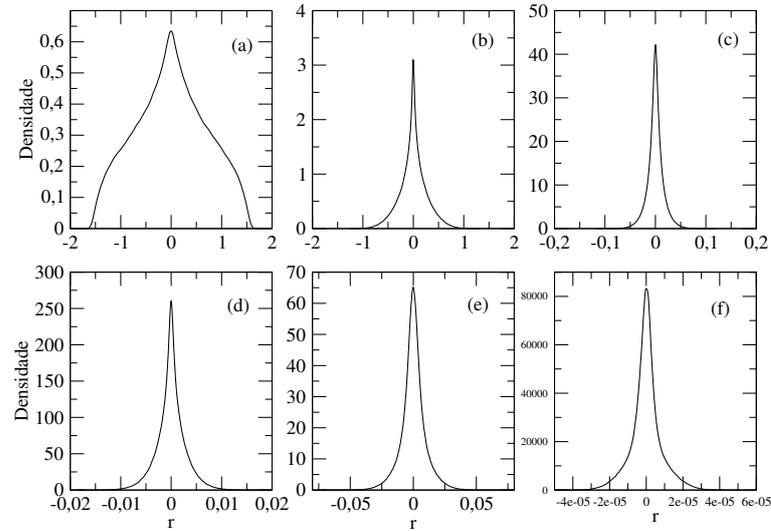
Figura 7.2: Retornos para as séries investigadas. (a) $a = 0,00$, (b) $a = 1,00$, (c) $a = 1,50$, (d) $a = 2,00$, (e) $a = 2,50$, (f) $a = 3,00$



Fonte: Autor, 2014.

Nas figuras 7.1 e 7.2 é possível, respectivamente, ver os dados originais, gerados pela equação 6.9 e os dados de retorno, que são calculados com a equação 4.2, enquanto que na figura 7.3 pode-se observar a densidade de probabilidade para os dados abordados, os retornos. Nenhuma das curvas tem características gaussianas, que é o primeiro passo para os testes. É possível inferir visualmente que as curvas são aparentemente leptúrticas, ou seja, uma *kurtosis* maior que a gaussiana.

Figura 7.3: Densidades de probabilidade (a) $a = 0,00$, (b) $a = 1,00$, (c) $a = 1,50$, (d) $a = 2,00$, (e) $a = 2,50$, (f) $a = 3,00$



Fonte: Autor, 2014.

Afim de determinar os parâmetros iniciais das séries, foram calculados os terceiros e quartos momentos para todas as séries presentes nas análises. Esses resultados podem ser observados nas tabelas 7.1 e 7.2, a seguir:

Tabela 7.1: Valores iniciais da *skewness* para as séries geradas do movimento browniano fracionado

$a =$	Valores iniciais da <i>skewness</i>
0.00	-0.00095
0.25	-0.00193
0.50	$7.17231 \cdot 10^{-05}$
0.75	0.00032
1.00	-0.00339
1.25	-0.00142
1.50	-0.00167
1.75	0.00250
2.00	0.00692
2.25	-0.01894
2.50	0.00021
2.75	0.01939
3.00	0.07243

Fonte: Autor, 2014.

Tabela 7.2: Valores iniciais da *kurtosis* para as séries geradas do movimento browniano fracionado

$a =$	Valores iniciais da <i>kurtosis</i>
0.00	-0.68472
0.25	-0.59428
0.50	-0.4071539
0.75	0.04358
1.00	1.08334
1.25	2.06849
1.50	2.45372
1.75	2.04917
2.00	2.02993
2.25	2.14443
2.50	2.44105
2.75	1.96785
3.00	1.95789

Fonte: Autor, 2014.

Claramente as séries não possuem um perfil gaussiano. Existe uma assimetria em torno do eixo central da distribuição, a *skewness*, bem sutil, e o quarto momento, a *kurtosis* excessiva, subtraída do valor gaussiano, é diferente de zero em todos os casos.

Uma condição muito importante e pertinente aos estudos aqui apresentados, é a condição de Lindeberg, comentada anteriormente. Apenas séries que satisfaçam a condição de Lindeberg, podem ser usadas na aplicação do TSL. Afim de verificar essa concordância, testamos os dados da série como requisita o teorema, e vemos que a convergência do parâmetro é muito rápida, o que indica que realmente temos séries que possuem a confiabilidade necessária para a aplicação do teorema.

7.1 Análise da Convergência

Será utilizada a mesma metodologia abordada no capítulo 5, onde foram realizados os testes para a convergência dos momentos estatísticos, *skewness* e *kurtosis*, além dos testes KS e KL, previamente discutidos.

Para gerar as séries segundo os teoremas, foram utilizadas as equações e determinações já exploradas. No caso do TLC, a agregação é realizada segundo a equação 3.1 e para o TSL a análise foi iniciada com uma seção de $t = 10^{-13}$ seguido de incrementos de $\Delta t = 2,5 \cdot 10^{-7}$, sendo suficientemente pequenos para garantir que existe uma variação suave nos dados.

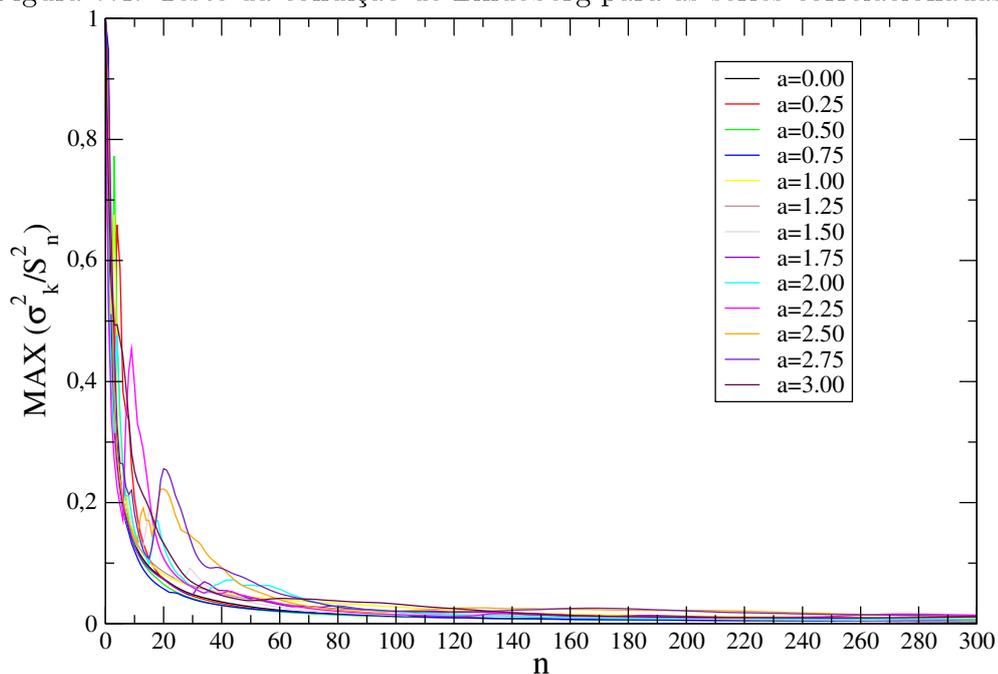
Assim como anteriormente, a tempo da da seção é determinado por:

$$\tau_t = \frac{\sigma_{S_t}^2}{\sigma_{S_n}^2}, \quad (7.1)$$

enquanto para o TLC, usa-se apenas o número de variáveis agregadas.

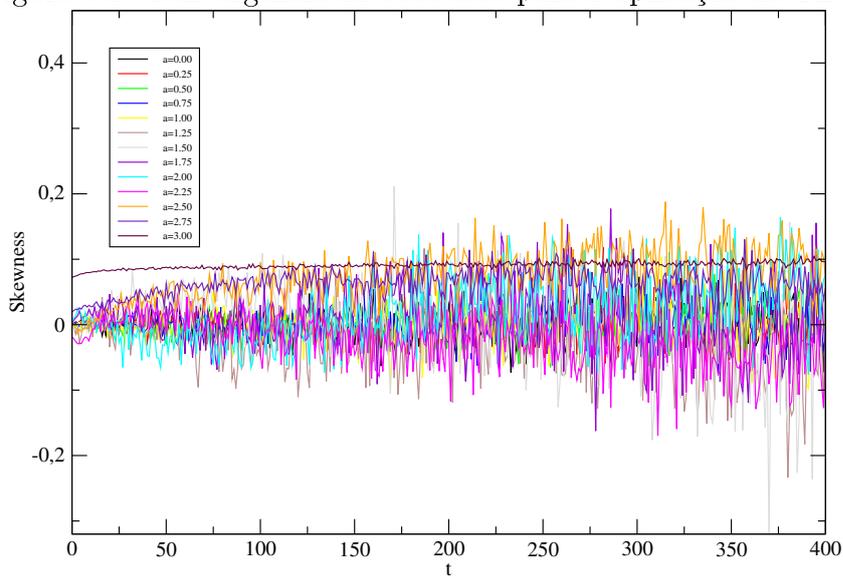
Como o TSL utiliza a variância para os incrementos, as séries mais próximas a gaussiana, ou seja, as menos correlacionadas, não se destoam muita da série original, tendo

Figura 7.4: Teste da condição de Lindeberg para as séries correlacionadas.



Fonte: Autor, 2014.

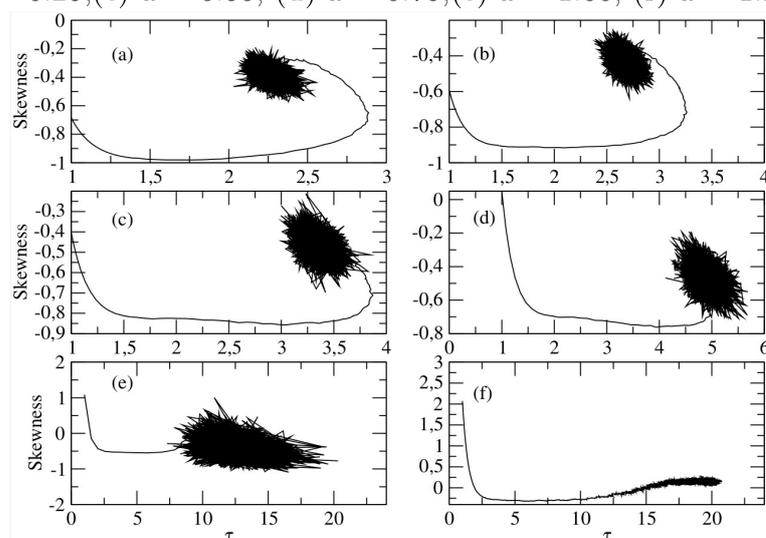
um tempo da seção limitante e criando alguns dados ruidosos mas que convergem rapidamente.

Figura 7.5: Convergência da *skewness* para a aplicação do TLC.

Fonte: Autor, 2014.

As figuras que retratam o perfil da *skewness* para a aplicação do TLC, que são 7.5, 7.6 e 7.7, apesar de algumas flutuações, é possível observar uma tendência do regime convergir para um valor próximo de 0, que corresponde ao valor gaussiano. Mas isso

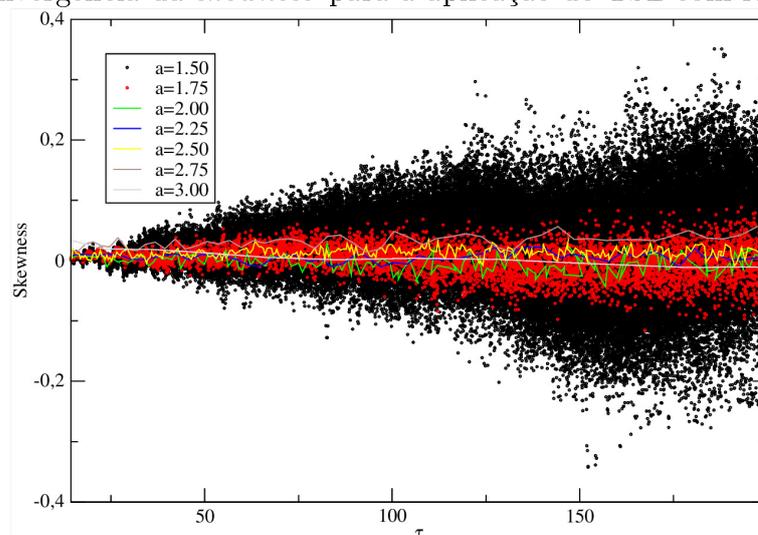
Figura 7.6: Convergência da *skewness* para a aplicação do TSL com baixa correlação. (a) $a = 0.00$, (b) $a = 0.25$, (c) $a = 0.50$, (d) $a = 0.75$, (e) $a = 1.00$, (f) $a = 1.25$,



Fonte: Autor, 2014.

não é surpresa, uma vez que na tabela 7.1 mostrada nesse mesmo capítulo, os valores da iniciais da *skewness* para as séries já demonstra tal tendência.

Figura 7.7: Convergência da *skewness* para a aplicação do TSL com fortes correlações.



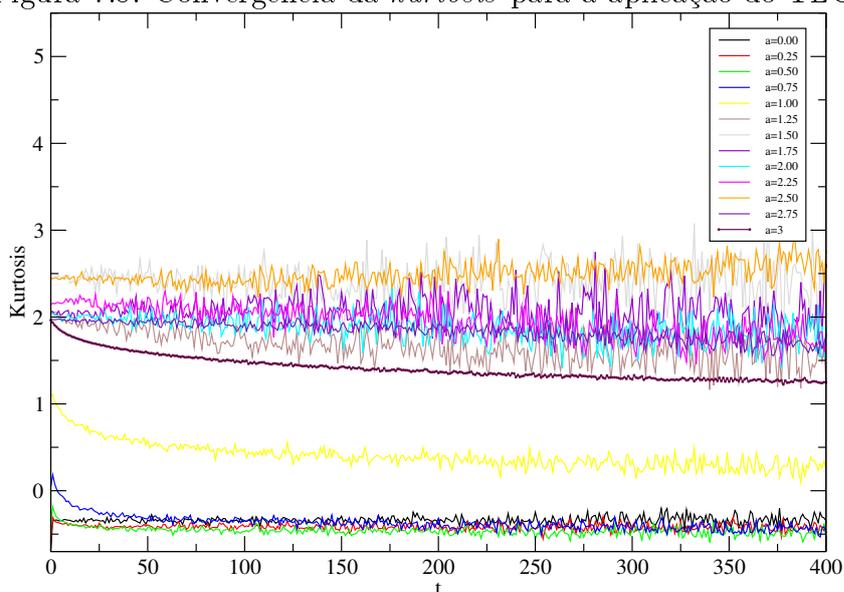
Fonte: Autor, 2014.

Para a *kurtosis* excessiva, exibida nas figuras 7.8, 7.9 e 7.10, onde o eixo está subtraído de 3, para a gaussiana ter o valor nulo, observa-se que nem todas as curvas seguem a tendência esperada para a convergência e esse regime e especialmente é observado para as curvas com correlação alta, o que indica que essa correlação impede efetivamente a ação do teorema.

Um outro detalhe que pode ser interpretado por esses dados, é o fato de que o TSL gera uma tendência dos dados a serem levados aos valores gaussianos, enquanto o TLC praticamente não altera os valores das séries, mostrando que o mesmo não age com efetividade sobre séries que são fortemente correlacionadas.

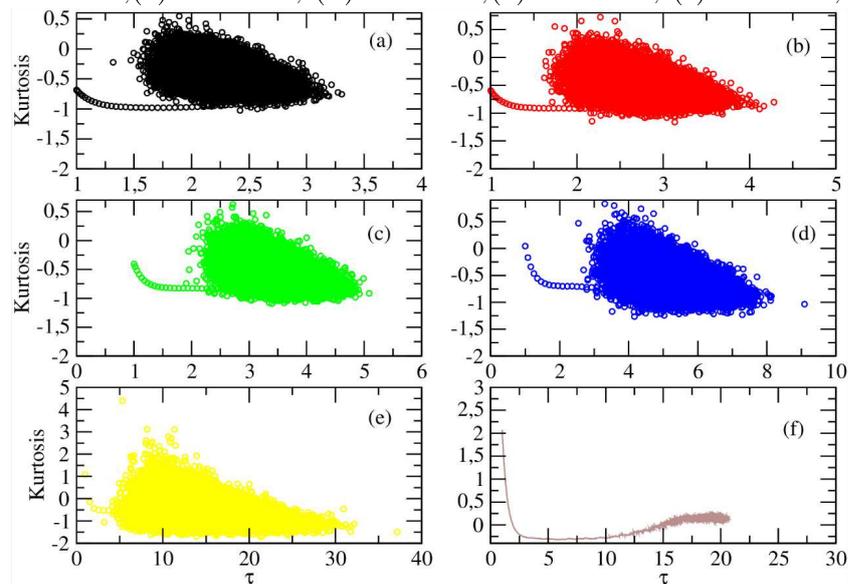
Usando o mesmo procedimento metodológico previamente abordado, foi aplicado o teste KS e o teste KL em todas as séries e os resultados se mostraram bastante intrigantes, tendo em vista que nenhum padrão é claramente observado ou descrito por meio dos gráficos nas figuras a seguir. A convergência para alguns testes e para algumas séries é clara, mas para outro teste é possível notar uma total disparidade entre o que é esperado e o que é visto no comportamento das curvas dos parâmetros analisados.

Figura 7.8: Convergência da *kurtosis* para a aplicação do TLC.



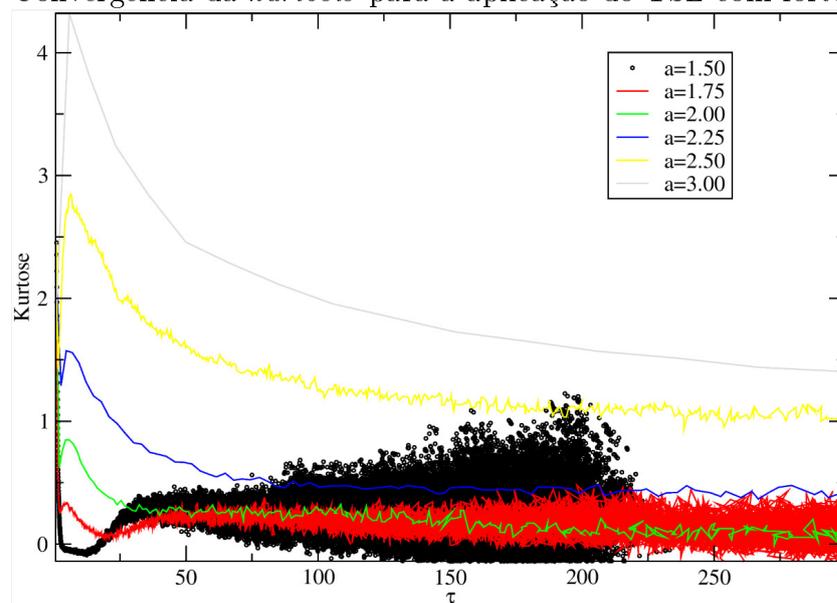
Fonte: Autor, 2014.

Figura 7.9: Convergência da *kurtosis* para a aplicação do TSL com baixa correlação. (a) $a = 0.00$, (b) $a = 0.25$, (c) $a = 0.50$, (d) $a = 0.75$, (e) $a = 1.00$, (f) $a = 1.25$,



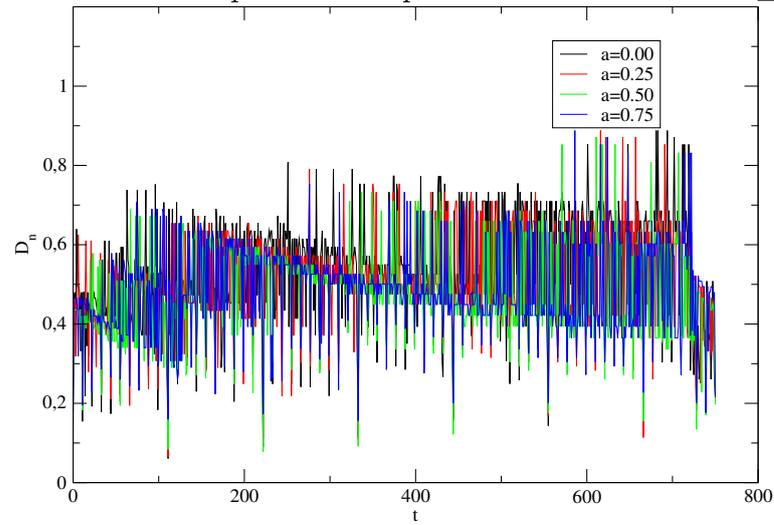
Fonte: Autor, 2014.

Figura 7.10: Convergência da *kurtosis* para a aplicação do TSL com fortes correlações.

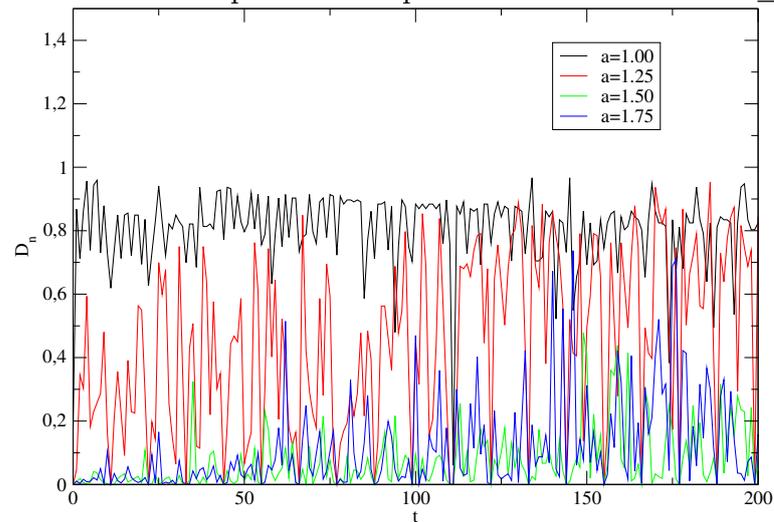


Fonte: Autor, 2014.

As figuras do teste KS, 7.11, 7.12, 7.13, 7.14 e 7.15, que é um teste não paramétrico que mede a maior distância que duas distribuições cumulativas de probabilidade possuem, mostra valores muito elevados para algumas séries e ruidosos, sem padrões ou leis aparentes. Sabe-se que o maior valor que a diferença poderá ser é exatamente 1, quando as distribuições são completamente destoantes. Esse resultado aparente choca com outras visões que indicavam a convergência das séries para determinados valores do parâmetro a .

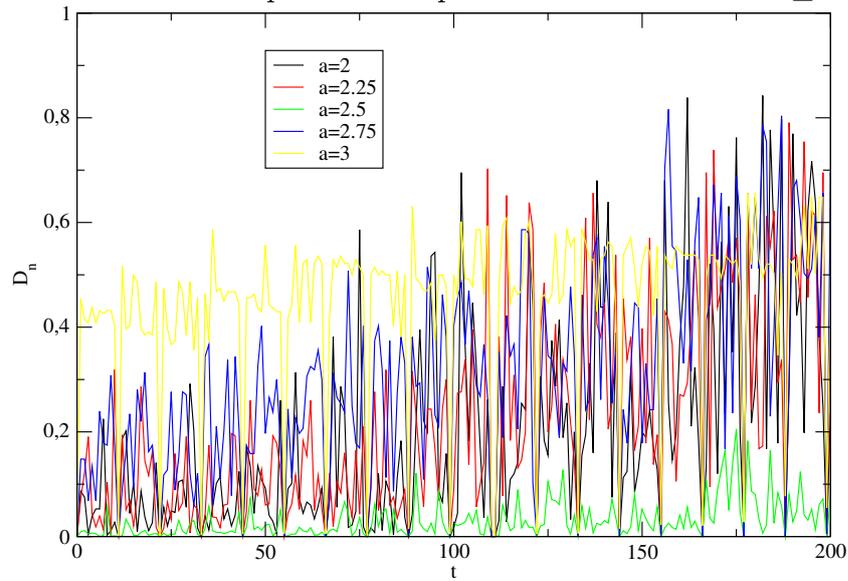
Figura 7.11: Teste KS para séries aplicadas ao TLC e com $0 \leq a < 1$.

Fonte: Autor, 2014.

Figura 7.12: Teste KS para séries aplicadas ao TLC e com $1 \leq a < 2$.

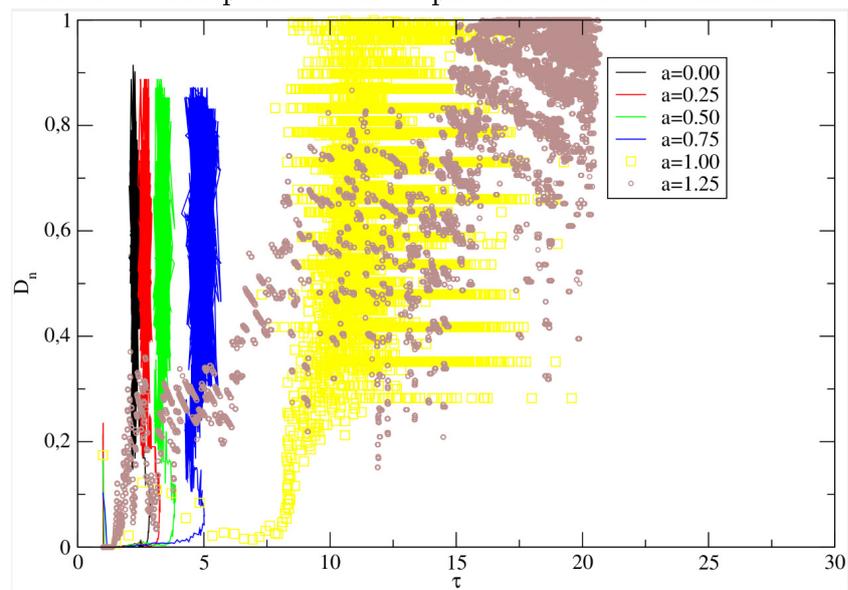
Fonte: Autor, 2014.

Por mais que os valores ainda não fossem gaussianos, não era de se esperar tais resultados.

Figura 7.13: Teste KS para séries aplicadas ao TLC e com $2 \leq a \leq 3$.

Fonte: Autor, 2014.

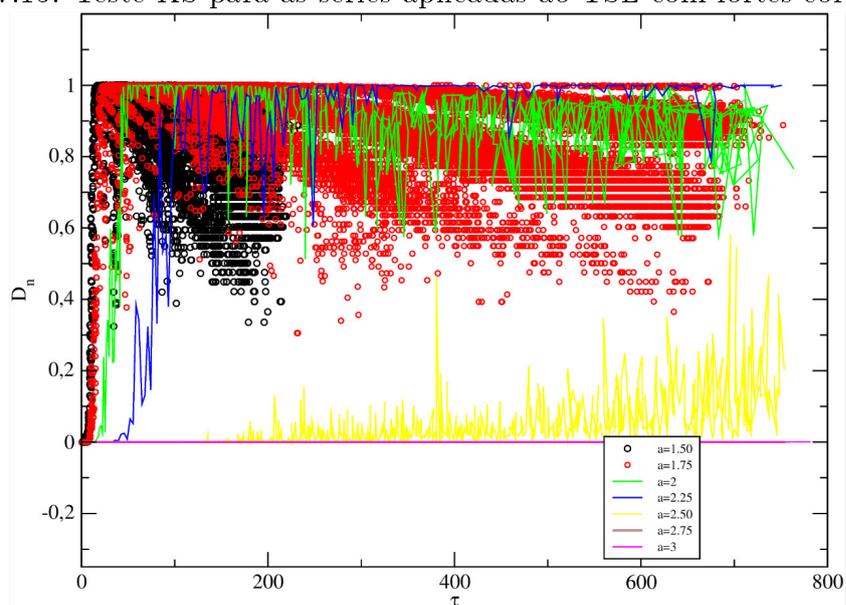
Figura 7.14: Teste KS para as séries aplicadas ao TSL com baixa correlação



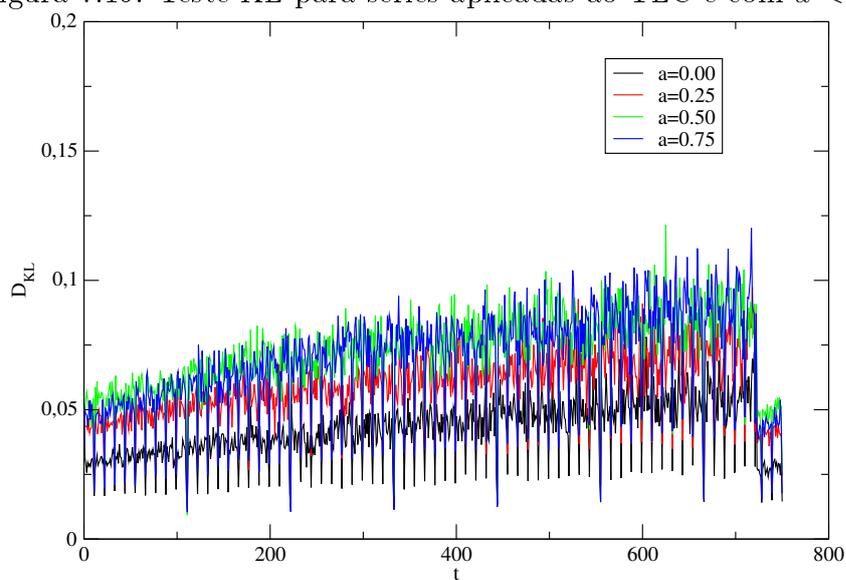
Fonte: Autor, 2014.

As figuras 7.16, 7.17, 7.18, 7.19 e 7.20, mostram o comportamento das séries, sob a aplicação do teste KL, que mede de forma indireta o grau de informação perdido, ao transformar uma distribuição em outra.

Figura 7.15: Teste KS para as séries aplicadas ao TSL com fortes correlações.

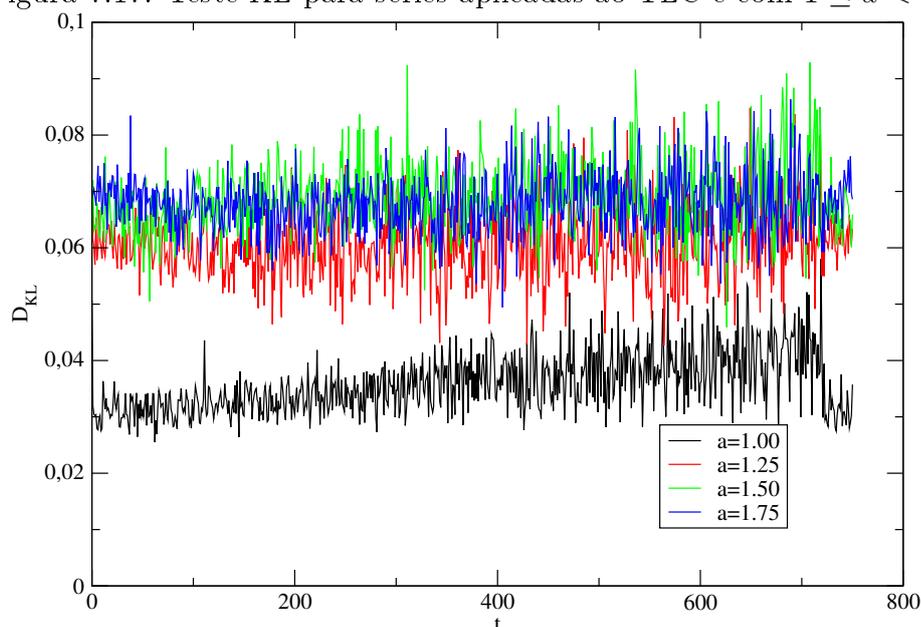


Fonte: Autor, 2014.

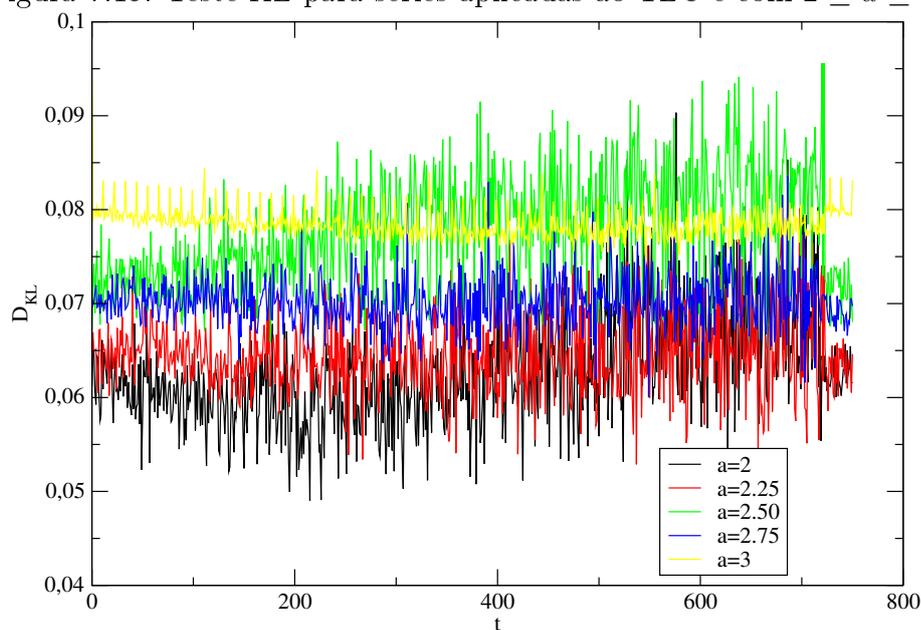
Figura 7.16: Teste KL para séries aplicadas ao TLC e com $a < 1$.

Fonte: Autor, 2014.

A despeito da figura 7.19, que mostra grandes flutuações no resultado e valores muito altos para o que era de se esperar, as outras curvas apresentam padrões já esperados

Figura 7.17: Teste KL para séries aplicadas ao TLC e com $1 \leq a < 2$.

Fonte: Autor, 2014.

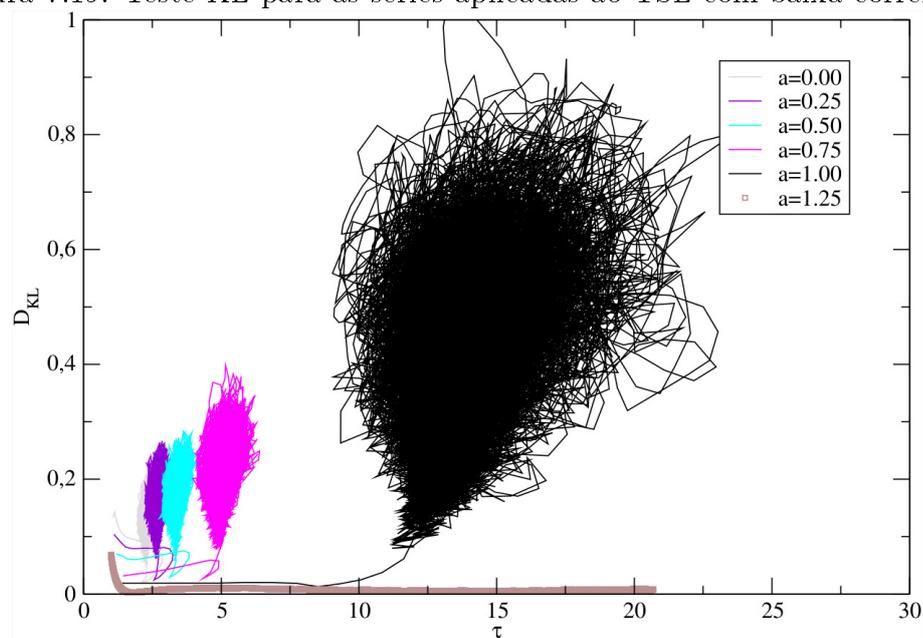
Figura 7.18: Teste KL para séries aplicadas ao TLC e com $2 \leq a \leq 3$.

Fonte: Autor, 2014.

de convergência, onde o parâmetro calculado pelo teste, D_{KL} tende a zero conforme a quantidade de termos é agregada no caso do TLC ou o tempo de seção cresce, no caso do TSL. Esse resultado entra em contradição com o teste KS, onde foi observado que as séries não convergiam.

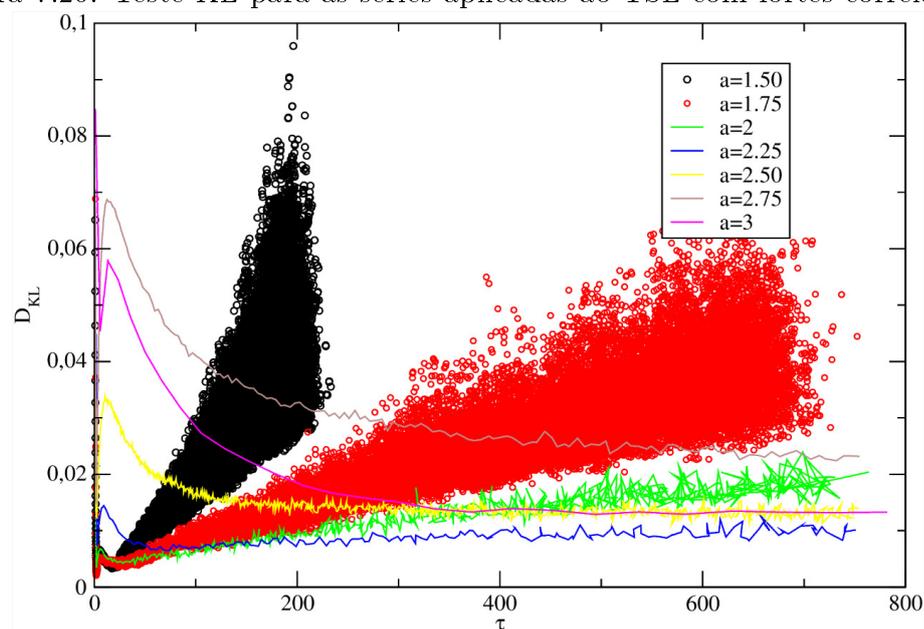
O ponto mais interessante a ser discutido, encontra-se exatamente na figura que não segue o padrão. A figura 7.19 mostra que apesar de possuir uma correlação muito baixa,

Figura 7.19: Teste KL para as séries aplicadas ao TSL com baixa correlação



Fonte: Autor, 2014.

Figura 7.20: Teste KL para as séries aplicadas ao TSL com fortes correlações.



Fonte: Autor, 2014.

os dados acabam se tornando ruidosos e sua análise estatística comprometida pelo TSL, enquanto isso, na figura 7.16, é clara a convergência, pois os valores ali mostrados sugerem que a informação é pouco perdida. Esse tipo de abordagem cria uma leve segregação entre os teoremas, o teorema de Lévy, deveria ser uma generalização, ou seja, apresentar os mesmos resultados que o teorema do limite central, quando temos um regime descor-

relacionado.

O fato do regime gaussiano não ter sido alcançado em nenhum dos dois sistemas, segundo as análises propostas em observar alguns dos momentos estatísticos e os testes KS e KL, levanta a questão da validade dos teoremas quando aplicados à séries finitas. Possivelmente o problema aqui observado, deve-se a limitação que as séries possuem, apesar de muito grandes, ambos os teoremas estão condicionados ao limite de infinitas variáveis, algo que nunca será observado em séries temporais reais.

Mesmo dentro desse ponto de vista, é importante observar que o regime de convergência apesar de falho em algumas instâncias, ainda apresenta um comportamento tendencial das séries, e portanto, talvez ainda exista algum parâmetro, além da finitude da série, que controle a convergência para séries reais.

Conclusões e perspectivas

No trabalho aqui apresentado foi possível verificar como o comportamento extremamente complexo do sistema dinâmico econômico apresenta uma riqueza de propriedades que são fundamentais para o estudo de aplicações e da generalidade do Teorema do Limite Central.

O Teorema do limite central afirma que a agregação de variáveis aleatórias independentes via uma soma direta, resulta em uma distribuição de probabilidades normal, ou gaussiana. O teorema das seções de Levy mostra-se como uma generalização do TLC, permitindo a convergência da distribuição para uma distribuição normal, utilizando como parâmetro de agregação a variância da série temporal.

Para analisar o processo de convergência, foram aplicados esses dois teoremas a uma série temporal com inúmeras propriedades, entre as quais, destacam-se, as correlações de longo alcance no módulo dos retornos, a multifractalidade da série e uma distribuição de probabilidades com cauda grossa e leptocúrtica.

A aplicação dos dois teoremas na série temporal de câmbio do Marco Alemão frente ao Dólar Americano, nos mostrou que o sistema passa de um regime de convergência para um regime estacionário nos diversos testes de comparação utilizados e tais valores onde a série se tornou estacionária, em relação a convergência para o comportamento gaussiano, em nenhum dos teoremas foi verificado uma convergência total para valores gaussianos.

Os testes aplicados foram as análises de alguns momentos estatísticos, como a skewness e a kurtose excessiva, onde não se atingiram os valores gaussianos da nulidade, mesmo em média. Aplicamos também alguns testes comparativos de extrema eficácia e os mesmos mostraram que apesar de valores bem próximos a zero, o comportamento não é claramente gaussiano.

Por fim da análise dos sistemas econofísicos foi feita uma abordagem ao aspecto multifractal do sistema e este se mostrou quase inalterado pela aplicação dos teoremas gerando assim um questionamento pertinente, a multifractalidade do sistema gera correlações de longo alcance tão seguramente fortes que torna a convergência para o regime gaussiano impossível mesmo com a aplicação do TSL? Ou ela cria isso pela finidade de números que temos na série?

A análise do confronto entre séries correlacionadas com comportamento multifractal e comportamento monofractal se faz muito relevante e interessante na perspectiva apresentada nesse trabalho.

Afim de investigar melhor as lacunas existentes na análise inicial, foram geradas diversas séries cujas correlações eram determinadas através de um movimento browniano

fracionado. Essas investigações levaram a novas perguntas, tanto sobre a convergência a gaussiana quanto sobre os parâmetros limitantes da aplicação do TSL, que mostrou que séries fracamente correlacionadas ou completamente descorrelacionadas, sua aplicação se torna confusa e infrutífera.

Nesses sistemas físicos a correlação é um parâmetro ajustável, sendo então de mais fácil análise, diferente de sistemas econômicos, onde a correlação é uma propriedade intrínseca ao sistema. Mesmo com esses padrões não foi possível extrair resultados físicos relevantes das análises, pois os sistemas não apresentaram uma relação clara entre o regime de convergência e a autocorrelação da série.

Por fim a tese mostra que os teoremas aqui aplicados carecem de mais análises e afirmações, pois a maneira como estão expostos e abordados nos diversos âmbitos da literatura não nos permite concluir sobre sua abrangência. É preciso investir em confrontar os dados com séries provenientes de várias aplicações e fundamentos, para que seja coerente concluir o formalismo para um teorema geral e amplo, que comporte séries finitas e correlacionadas.

Nossos trabalhos foram publicados na revista *Europhysics Letters* na referência [16].

Formalismo matemático da condição de Lindeberg

A condição de Lindeberg[72] é uma condição suficiente para que o teorema do limite central seja válido para uma sequência de variáveis aleatórias independentes e apenas exige que tenham média e variância finitas.

Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidades, e $X_k : \Omega \rightarrow \mathbb{R}$, $k \in \mathbb{N}$, uma variável aleatória independente que pertence a esse espaço. Suponha que os valores esperados $\mathbb{E}X_k = \mu_k$ e as variâncias $\text{Var} X_k = \sigma_k^2$ existam e são finitas. Vamos definir também $s_n^2 := \sum_{k=1}^n \sigma_k^2$.

A condição de Lindeberg para a sequência de variáveis aleatórias X_k e tal que:

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{k=1}^n \int_{\{|X_k - \mu_k| > \varepsilon s_n\}} (X_k - \mu_k)^2 d\mathbb{P} = 0, \quad \text{para todo } \varepsilon > 0,$$

(onde essa integral é uma integral de Lebesgue sobre o conjunto $\{|X_k - \mu_k| > \varepsilon s_n\}$)

a integral de Lebesgue é uma generalização da integral de Riemann. Originalmente definida para funções $f : \mathbb{R}^n \rightarrow \mathbb{R}$, a integral de Lebesgue apresenta diversas vantagens em relação à integral de Riemann sobretudo em relação a processos de limite. Para maiores informações sobre a integral de Lebesgue veja Dudley, Richard M. (1989). Real analysis and probability, que trata da integração de Lebesgue no âmbito de probabilidades com ótimas referências históricas.

A condição de Lindeberg implica que

$$\max_{k=1, \dots, n} \frac{\sigma_k^2}{s_n^2} \rightarrow 0$$

com $n \rightarrow \infty$, isso garante que a contribuição de cada uma das variáveis que compõem X_k vão contribuir minimamente para a variância s_n^2 para valores suficientemente grandes de n . Nesse contexto a condição de Lindeberg nos fornece as bases onde o TCL se estabelece.[73]

Referências

- [1] Encyclopedia britânica disponível online em: <http://www.britannica.com/topic/gambling>.
- [2] CARDANO, G. *The book on games of chance (liber de ludo aleae)*. Rinehart and Winston, 1961.
- [3] BELLHOUSE, D. Decoding cardano's liber de ludo aleae. *Historia Mathematica*, v. 32, p. 180–202, 2005.
- [4] STEWART, I. *17 equações que mudaram o mundo*. Zahar, 2013.
- [5] BERNOULLI, J. *The art of conjecturing*. JHU Press, 2006.
- [6] SMITH, D. E. *A source book in mathematics*. Courier Dover Publications, 1959.
- [7] DE MOIVRE, A. *Doctrine of chances*. Disponível gratuitamente em <http://books.google.com.br>, 1738.
- [8] QUETELET, A. *A treatise on man and the development of his faculties*. Edinburgh : W. and R. Chambers, 1842.
- [9] PARETO, V. *Cours déconimie politique*. Université de Lausanne, 1896.
- [10] GAUSS, C. F. *Theory of the motion of the heavenly bodies moving about the sun in conic sections*. Boston: Little, Brown and Company; disponível gratuitamente em <https://archive.org/details/motionofheavenly00gausrich>, 1857.
- [11] FISCHER, H. *A history of the central limit theorem: from classical to modern probability theory*. Springer, 2011.
- [12] PóLYA, G. *Collected papers*. Cambridge, Cambridge.
- [13] MAXWELL, J. Illustrations of the dynamical theory of gases.—part i. on the motions and collisions of perfectly elastic spheres. *Philosophical Magazine*, v. 19, p. 19–32, 1860.
- [14] NICOLIS, G.; NICOLIS, C. Foundations of complex systems. *European Review*, v. 17, n. 2, p. 237–248, 2009.
- [15] LéVY, P. *Calcul des probabilités*. Gauthier-Villars, 1925.

- [16] NASCIMENTO, C. M.; HELENA, E.; PASSOS, F.; GLERIA, I.; FIGUEIREDO, A.; VISWANATHAN, G. Lévy sections vs. partial sums of heteroscedastic time series. *Europhysics Letters*, Les Ulis, v. 96, p. 68400, 2011.
- [17] EINSTEIN, A. *Investigations on the theory of the brownian motion*. Dover Publications, 1926, reimpresso em 1956.
- [18] ARFKEN, G.; WEBER, H. *Mathematical methods for physicists*. Elsevier, 2007.
- [19] NOLAN, J. P. *Stable distributions - models for heavy tailed data*. Birkhauser, Em progresso, Capítulo 1 online em <http://academic2.american.edu/jpnolan/stable/stable.html>, 2011.
- [20] O programa pode ser gratuitamente baixado em <http://www.r-project.org/>.
- [21] GLERIA, I.; MATSUSHITA, R.; SILVA, S. D. Sistemas complexos, criticalidade e leis de potencia. *Revista Brasileira de Ensino de Física*, v. 26, p. 99, 2004.
- [22] BARABÁSI, A.-L.; STANLEY, H. *Fractal concepts in surface growth*. Cambridge University Press, 1995.
- [23] STANLEY, H.; MEAKIN, P. Multifractal phenomena in physics and chemistry. *Nature*, London, v. 335, p. 405–409, 1988.
- [24] VISWANATHAN, G.; DA LUZ, M. G. E.; RAPOSO, P. E.; STANLEY, H. E. *The physics of foraging*. Cambridge University Press, 2011.
- [25] GOLDBERGER, A. L. Non-linear dynamics for clinicians: chaos theory, fractals, and complexity at the bedside. *Lancet*, v. 347, p. 1312–1314, 1996.
- [26] MANDELBROT, B. B. *The fractal geometry of nature*. Freeman, 1982.
- [27] FEDER, J. *Fractals*. Plenum Press, 1988.
- [28] HURST, H. E. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, v. 116, p. 770–799, 1951.
- [29] MANDELBROT, B. B.; WALLIS, J. R. Robustness of the rescaled range r/s in the measurement of noncyclic long run statistical dependence. *Water Resources Research*, v. 5, p. 967–988, 1969.
- [30] ARGOLO, C.; MARIZ, A.; MIYAZIMA, S. Correlation analysis of the damage spreading problem in a 2-dimensional ising model. *Physica A: Statistical Mechanics and its Applications*, v. 264, p. 142–148, 1999.
- [31] PENG, C.-K.; BULDYREV, S. V.; HAVLIN, S.; SIMONS, M.; STANLEY, H. E.; GOLDBERGER, A. L. Mosaic organization of dna nucleotides. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, v. 49, p. 1685–1689, 1994.
- [32] STANLEY, H. E. *Introduction to phase transitions and critical phenomena*. Oxford University Press, 1971.

- [33] HALSEY, T. C.; JENSEN, M. H.; KADANOFF, L. P.; PROCACCIA, I.; SHRAIMAN, B. I. Fractal measures and their singularities: The characterization of strange sets. *Physical Review A*, Woodbury, v. 33, p. 1141–1151, 1986.
- [34] OTT, E. *Chaos in dynamical systems*. Cambridge University Press, 1993.
- [35] CHHABRA, A.; JENSEN, R. V. Direct determination of the $f(\alpha)$ singularity spectrum. *Physical Review Letters*, Woodbury, v. 62, p. 1327–1330, Mar 1989.
- [36] KANTELHARDT, J. W.; ZSCHIEGNER, S. A.; KOSCIELNY-BUNDE, E.; HAVLIN, S.; BUNDE, A.; STANLEY, H. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, v. 316, p. 87 – 114, 2002.
- [37] LÉVY, P. Propriétés asymptotiques des sommes des variables aléatoires enchainées. *Bulletin de la Societé mathématique de France*, v. 59, p. 1–32, 1935.
- [38] LÉVY, P. *Théorie de l'addition des variables aléatoires*. Gauthier–Villars, 1937.
- [39] FIGUEIREDO, A.; GLERIA, I.; MATSUSHITA, R.; SILVA, S. D. Levy sections theorem revisited. *Journal of Physics A: Mathematical and Theoretical*, v. 40, p. 5783, 2007.
- [40] NASCIMENTO, C. M.; JÚNIOR, H. B. N.; JENNINGS, H. D.; SERVA, M.; GLERIA, I.; VISWANATHAN, G. M. Multifractality and heteroscedastic dynamics: An application to time series analysis. *EPL (Europhysics Letters)*, v. 81, p. 18002, 2008.
- [41] NETO, A.; MATSUSHITA, R.; SILVA, S.; SERVA, M.; VISWANATHAN, G. M.; NASCIMENTO, C. M.; GLERIA, I. M. The lévy sections theorem: An application to econophysics. *Physica A: Statistical Mechanics and its Applications*, v. 386, p. 756–759, 2007.
- [42] KOLMOGOROV, A. N. Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, v. 4, p. 83–91, 1933.
- [43] CONOVER, W. J. *Practical nonparametric statistics*. Wiley, 1999.
- [44] GLIVENKO, V.; CANTELLI, F. P. Sulla determinazione empirica delle leggi di probabilita. *Giornale dell'Istituto Italiano degli Attuari*, v. 4, p. 92–99, 1933.
- [45] KULLBACK, S.; LEIBLER, R. A. On information and sufficiency. *Annals of Mathematical Statistics*, v. 22, p. 79–86, 1951.
- [46] SHANNON, C. E. A mathematical theory of communication. *The Bell System Technical Journal*, v. 27, p. 379–423, 623–656, 1948.
- [47] BACHELIER, L. *Théorie de la spéculation*. 1900. Tese (Doutorado em Física) - Université de Paris, 1900.
- [48] FELLER, W. *An introduction to probability theory and its applications. vol i*. 1950.
- [49] MANDELBROT, B. B. The variation of certain speculative prices. *Journal of Business*, v. 36, p. 394–419, 1963.

- [50] BLACK, F.; SCHOLES, M. The pricing of options and corporate liabilities. *The Journal of Political Economy*, v. 81, p. 37–654, 1973.
- [51] KADANOFF, L. From simulation model to public policy: An examination of forrester's urban dynamics. *Simulation*, v. 16, p. 261–268, 1971.
- [52] MONTROLL, E.; BADGER, W. *Introduction to quantitative aspects of social phenomena*. Gordon and Breach, 1974.
- [53] MANTEGNA, R.; STANLEY, H. Scaling behaviour in the dynamics of an economic index. *Nature*, London, v. 376, p. 46–49, 1995.
- [54] MANTEGNA, R.; STANLEY, H. *An introduction to econophysics: Correlations and complexity in finance*. Cambridge University Press, 2000.
- [55] FREIRE, R. R.; MORICONI, L.; RIBEIRO, M. B. Encontro de econofísica 2007, 9 de Novembro de 2007.
- [56] FREIRE, R. R.; MORICONI, L.; RIBEIRO, M. B.; ROSENFELD, R. Econofis' 10, 2010.
- [57] QUEIRÓS, S. D.; ANTENEODO, C.; CAJUEIRO, D. Econofis' 14, 2014.
- [58] BAR-YAM, Y. *Dynamics of complex system*. Perseus, Reading, Mass., 1997.
- [59] FERREIRA, A. S.; CRESSONI, J. C.; VISWANATHAN, G. M.; ALVES DA SILVA, M. A. Anomalous diffusion in non-markovian walks having amnestically induced persistence. *Physical Review E*, v. 81, p. 011125, 2010.
- [60] HORST, U. Financial price fluctuations in a stock market model with many interacting agents. *Economic Theory*, v. 25, p. 917–932, 2005.
- [61] F.S.PASSOS; NASCIMENTO, C.; GLERIA, I.; DA SILVA, S.; VISWANATHAN, G. M. Fat tails, long-range correlations and multifractality as emergent properties in nonstationary time series. *Europh*, v. 93, p. 58006, 2011.
- [62] BROWN, R. *The miscellaneous botanical works of robert brown volume 1*. Robert Hardwicke, 1866.
- [63] MANDELROT, B. B.; NESS, J. W. V. Fractional brownian motions, fractional noises and applications. *SIAM review*, v. 10, p. 422–437, 1968.
- [64] LÉVY, P. Random functions: general theory with special reference to laplacian random functions. *University of California Publications in Statistics*, v. 1, p. 331–390, 1953.
- [65] DE MOURA, F. A. B. F.; LYRA, M. L. Delocalization in the 1d anderson model with long-range correlated disorder. *Physical Review Letters*, Woodbury, v. 81, p. 3735–3738, 1998.
- [66] ANDERSON, P. W. Absence of diffusion in certain random lattices. *Physical Review*, New York, v. 109, p. 1492–1505, 1958.

- [67] GREIS, N. P.; GREENSIDE, H. S. Implication of a power-law power-spectrum for self-affinity. *Physical Review A*, Woodbury, v. 44, p. 2324–2334, 1991.
- [68] HIGUCHI, T. Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena*, v. 31, p. 277 – 283, 1988.
- [69] FOX, C. G. Empirically derived relationships between fractal dimension and power law form frequency spectra. *Pure and Applied Geophysics*, v. 131, p. 211–239.
- [70] OSBORNE, A.; PROVENZALE, A. Finite correlation dimension for stochastic systems with power-law spectra. *Physica D: Nonlinear Phenomena*, v. 35, p. 357 – 381, 1989.
- [71] HIGUCHI, T. Relationship between the fractal dimension and the power law index for a time series: A numerical investigation. *Physica D: Nonlinear Phenomena*, v. 46, p. 254 – 264, 1990.
- [72] LINDEBERG, J. W. Eine neue herleitung des exponentialgesetzes in der wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, v. 15, p. 211–225.
- [73] Parte do texto foi retirado da pagina http://en.wikipedia.org/wiki/lindeberg's_condition.