
UNIVERSIDADE ESTADUAL DE MARINGÁ
DEPARTAMENTO DE FISICA

Tese de Doutorado

**Física estatística dos sistemas complexos:
Aplicações interdisciplinares**

Sergio de Picoli Junior

Orientador: Prof. Dr. Renio dos Santos Mendes

2007

UNIVERSIDADE ESTADUAL DE MARINGÁ
DEPARTAMENTO DE FISICA

**Física estatística dos sistemas complexos:
Aplicações interdisciplinares**

Sergio de Picoli Junior

Tese de Doutorado
submetida ao Departamento de Física da
Universidade Estadual de Maringá

Orientador:

Prof. Dr. Renio dos Santos Mendes

Banca Examinadora:

Prof. Dr. José Soares Andrade Junior - UFC/CE

Prof. Dr. Luciano Rodrigues da Silva - UFRN/RN

Prof. Dr. Thadeu Josino Pereira Penna - UFF/RJ

Prof. Dr. Ervin Kaminski Lenzi - UEM/PR

Prof. Dr. Luiz Roberto Evangelista - UEM/PR

Prof. Dr. Mauro Luciano Baesso - UEM/PR

2007

“I have come to know that everything that the true God makes, it will prove to be to time indefinite. To it there is nothing to add and from it there is nothing to subtract... I have seen the occupation that God has given to the sons of mankind in which to be occupied... Even time indefinite he has put in their heart, that mankind may never find out the work that the true God has made from the start to the finish. ”

— Ecclesiastes 3:10,11,14

“Vim saber que tudo o que o verdadeiro Deus faz mostrará ser por tempo indefinido. Não há nada a acrescentar-lhe e não há nada a subtrair-lhe... Vi a ocupação que Deus deu aos filhos da humanidade para se ocuparem nela... Pôs até mesmo tempo indefinido no seu coração, para que a humanidade nunca descobrisse o trabalho que o verdadeiro Deus tem feito do começo ao fim.”

— Eclesiastes 3:10,11,14

Agradecimentos

Agradeço especialmente ao professor Renio, orientador e amigo, pelas boas idéias, sugestões práticas e pelo interesse constante no desenvolvimento do trabalho. Agradeço também ao professor Luis Malacarne, apoiador incansável e participante ativo nas atividades de pesquisa desenvolvidas. De fato, é como a antigo provérbio já dizia: “Há frustração de planos quando não há palestra confidencial, mas na multidão de conselheiros há consecução”. (Provérbios 15:22)

Não poderia deixar de agradecer a todos aqueles que me apoiaram durante estes anos, principalmente meus pais, demais membros achegados da família, Juliana, amigos e colegas de curso.

Agradeço também ao Departamento de Física da Universidade Estadual de Maringá pela infra-estrutura e ao CNPq pelo apoio financeiro.

Por fim, faço minhas as seguintes palavras dirigidas ao criador: “Digno és, Jeová, sim, nosso Deus, de receber a glória, e a honra, e o poder, porque criaste todas as coisas e porque elas existiram e foram criadas por tua vontade.” (Revelação 4:11)

O presente trabalho é o resultado de anos de estudo e pesquisa, de erros e acertos... Olhando para a sua forma final, tive que resistir à idéia de reescrever os capítulos para torná-los mais claros e precisos. Espero que os leitores compreendam isso ao notarem estas e outras imperfeições.

Abstract

In the present work, we apply concepts and methods from statistical physics to study complex systems. Specifically, we analyze the dynamics of *i*) circulation of magazines; *ii*) citations to scientific journals; *iii*) religious activities of Jehovah's Witnesses; and *iv*) a geomagnetic signal - the Dst index. In all the cases we find evidence of scaling behavior and universality. For the systems *i*, *ii* and *iii*, we show that: 1) the distribution of sizes exhibits power law asymptotic behavior; 2) the distribution of growth rates displays a universal form which does not depend on the size or on the measure of size used; 3) the standard deviation of growth rates decays with size as a power law; 4) these findings are remarkably similar to those observed in the dynamics of economic and scientific research activities. Furthermore, we propose two stochastic models in order to reproduce distinct aspects of the observed behavior. Moreover, for the system *iv*, we show that: 1) the distribution of Dst increments exhibits the same functional form for a wide range of time scales, presenting power law asymptotic behavior; 2) Dst series are long-range correlated, being well approximated by a $1/f$ noise; 3) magnitude series of Dst increments are long-range correlated (persistent behavior); 4) sign series of Dst increments are long-range anti-correlated (anti-persistent behavior); 5) Dst series present multifractal properties; 6) periods of high and low geomagnetic activity exhibit similar multifractal properties; 7) these findings are surprisingly similar to those observed in the dynamics of a physiologic signal. In general, our findings may contribute to a better understanding of the mechanisms which govern the dynamics of the systems considered, serving as a guide to develop models reproducing the observed behavior.

Resumo

Neste trabalho, aplicamos conceitos e métodos de física estatística no estudo de sistemas complexos. Especificamente, analisamos a dinâmica *i*) da circulação de revistas; *ii*) das citações de periódicos científicos; *iii*) de atividades religiosas das Testemunhas de Jeová; e *iv*) de um sinal geomagnético - o índice Dst. Em todos os casos, encontramos evidências de invariância de escala e comportamento universal. Para os sistemas *i*, *ii* e *iii*, mostramos que 1) a distribuição dos tamanhos exibe comportamento assintótico do tipo lei de potência; 2) a distribuição das taxas de crescimento exibe uma forma universal que não depende do tamanho nem da medida de tamanho considerada; 3) o desvio padrão das taxas de crescimento decresce com o tamanho como lei de potência; 4) estes resultados são notavelmente similares aos observados na dinâmica de atividades econômicas e de pesquisa científica. Além disso, propomos dois modelos estocásticos que reproduzem aspectos distintos do comportamento observado. Ademais, para o sistema *iv*, mostramos que 1) a forma funcional da distribuição dos incrementos se mantém inalterada para diferentes escalas de tempo, apresentando comportamento assintótico do tipo lei de potência; 2) o índice Dst exibe correlações de longo alcance do tipo ruído $1/f$; 3) a série das magnitudes dos incrementos apresenta correlações de longo alcance (comportamento persistente); 4) a série dos sinais dos incrementos apresenta anticorrelações de longo alcance (comportamento anti-persistente); 5) o índice Dst exibe comportamento multifractal; 6) períodos de alta e baixa atividade geomagnética são caracterizados por propriedades multifractais similares; 7) estes resultados são surpreendentemente similares aos observados na dinâmica de um sinal fisiológico. De modo geral, os resultados descritos aqui podem contribuir para um melhor entendimento dos mecanismos que governam a dinâmica dos sistemas considerados, servindo de guia no desenvolvimento de modelos que reproduzam o comportamento observado.

Conteúdo

Abstract	6
Resumo	7
1 Introdução	10
2 Conceitos e métodos	14
2.1 Densidade de probabilidade	14
2.1.1 Histogramas e distribuições acumuladas	15
2.2 Distribuições do tipo lei de potência	16
2.3 Correlações de longo alcance	16
2.3.1 O método DFA	18
2.3.2 Detectando correlações com DFA	20
2.4 Multifractalidade	23
2.4.1 O método mf-DFA	24
2.4.2 Detectando multifractalidade com mf-DFA	25
3 Física estatística dos sistemas complexos: recentes aplicações	28
3.1 Dinâmica de crescimento de organizações	28
3.1.1 Atividades econômicas	29
3.1.2 Atividades de pesquisa científica	31
3.1.3 Características universais	33
3.1.4 Modelos estocásticos	34
3.2 Dinâmica dos sinais complexos	35
3.2.1 Sinal fisiológico	35
3.2.2 Sistema hidrodinâmico	40
3.2.3 Tráfego na Internet	42

4	Circulação de revistas e jornais	44
4.1	Introdução	44
4.2	Distribuição da circulação	46
4.3	Distribuição das taxas de crescimento	48
4.4	Características universais	50
4.5	Modelo estocástico I	51
5	Citações de periódicos científicos	55
5.1	Introdução	55
5.2	Distribuição do fator de impacto	57
5.3	Distribuição das taxas de crescimento	59
5.4	Características universais	62
5.5	Modelo estocástico II	64
6	Atividades religiosas das Testemunhas de Jeová	68
6.1	Introdução	68
6.2	Distribuição do número de membros e de suas taxas de crescimento	71
6.3	Tempo gasto, cursos dirigidos e pioneiros	75
6.4	Aspectos universais	78
6.5	Comparação com o modelo estocástico II	80
7	Sinal geomagnético: o índice Dst	83
7.1	Introdução	83
7.2	Distribuição das flutuações no índice Dst	85
7.3	Correlações de longo alcance	87
7.4	Multifractalidade	91
7.5	Características universais	93
8	Conclusão	95
A	Exemplos de distribuições	98
A.1	Distribuições Gaussiana e de Laplace	98
A.2	Leis de potência assintóticas	99
A.2.1	Distribuição q -exponencial	100
A.2.2	Distribuição q -Weibull	101
A.2.3	Distribuição q -Gaussiana	102
A.2.4	Distribuição simétrica de Lévy	103

Capítulo 1

Introdução

Por muitos anos, a especialização profissional tem levado a um isolamento progressivo de disciplinas individuais. Por outro lado, em anos recentes, tem sido possível unificar em uma única disciplina, campos de pesquisa bem definidos, tais como: física, química, biologia, geologia, lingüística, medicina, economia e ciências sociais. Trata-se do emergente campo de pesquisa dos “sistemas complexos” [1-9].

O dicionário Aurélio define a palavra “complexo” como “que abrange ou encerra muitos elementos ou partes” e “grupo ou conjunto de coisas, fatos ou circunstâncias que têm qualquer ligação ou nexos entre si”. No contexto deste emergente campo de pesquisa interdisciplinar, a expressão “sistema complexo” assume um significado ainda mais amplo. Tipicamente, sistemas complexos exibem características tais como: *i*) o sistema é composto de muitos elementos ou partes que interagem entre si; *ii*) os elementos e suas interações são heterogêneos, com características individuais distintas; *iii*) os elementos estão organizados em grupos ou hierarquias que também podem interagir entre si; *iv*) o sistema adapta-se diante de novas situações, à medida que evolui no tempo; e *iv*) o sistema exibe padrões globais que emergem das interações entre elementos individuais em diferentes escalas de observação (veja fig. 1.1).

Nesse amplo contexto, a física estatística tem encontrado interessantes e desafiadoras aplicações que vão muito além das fronteiras da física. A física estatística fornece uma visão global do sistema em estudo, sem a necessidade de uma descrição detalhada do mesmo. Este tipo de abordagem tem contribuído para mostrar como fenômenos e processos em diferentes áreas de pesquisa, por muito tempo considerados não relacionados entre si, podem ter uma descrição comum. De fato, conceitos e métodos de física estatística estão sendo aplicados no estudo dos mais diversos e intrigantes siste-

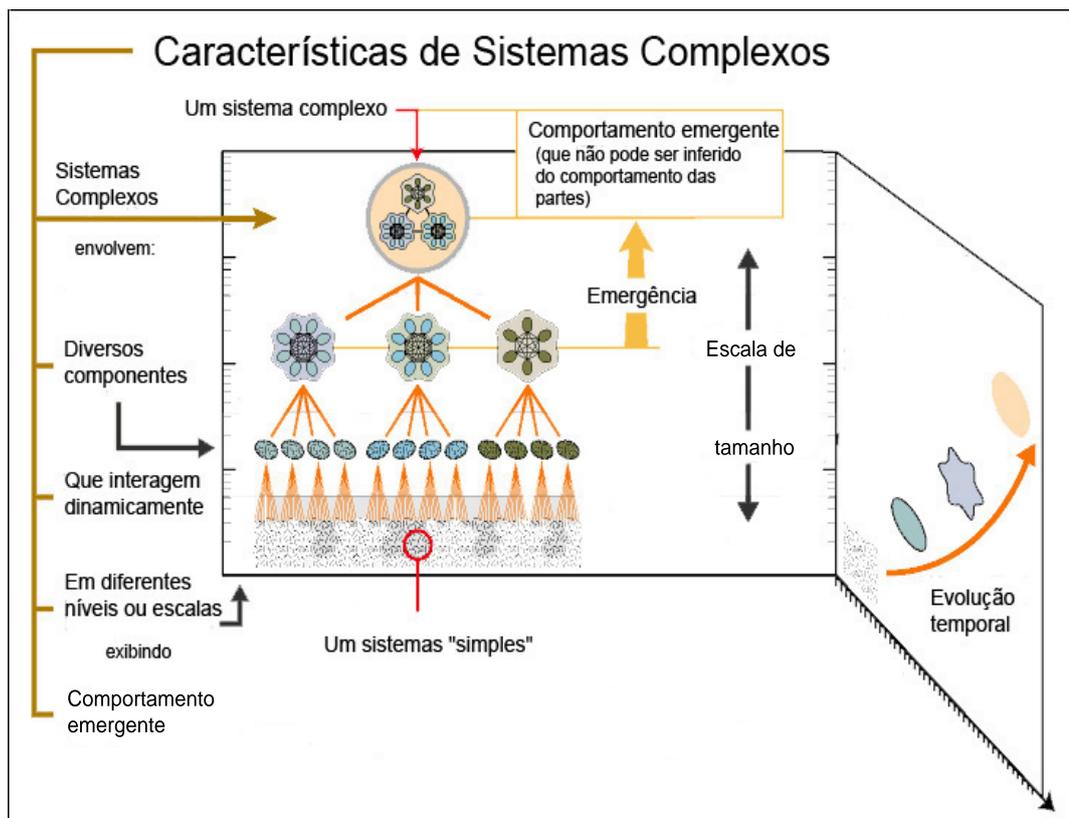


Figura 1.1: Diagrama de características típicas de sistemas complexos.

mas complexos - de sequências de DNA[10] a temperatura atmosférica[11]; de sinais fisiológicos[12] a índices econômicos[13, 14]; de populações de pássaros[15] a populações humanas[16, 17].

Tipicamente, os sistemas complexos são caracterizados por leis que exibem invariância de escala[18, 19]. Esta característica marcante se manifesta, por exemplo, quando a lei que descreve alguma variável de interesse no sistema mantém a sua forma funcional para diferentes escalas de observação. Nesse caso, uma simples redefinição de variáveis faz com que curvas distintas, relacionadas a diferentes escalas de observação, sejam agrupadas sob uma só curva. Nesse sentido, o sistema exibe invariância de escala quando alguma de suas variáveis de interesse exibe comportamento do tipo lei de potência.

Um outro conceito de importância fundamental na física estatística dos sistemas complexos refere-se à universalidade[18, 19]. Esta propriedade relaciona-se com a tendência de certos sistemas distintos exibirem leis empíricas similares. Os sistemas poderiam então ser divididos em grupos com a propriedade de que todos os sistemas dentro de um mesmo grupo exibem características em comum - sugerindo que os mecanismos responsáveis pelo comportamento observado também sejam similares. De fato, tais similaridades podem indicar a presença de mecanismos “universais” que atuam independentemente das particularidades do sistema em estudo. Este princípio de universalidade em sistemas complexos tem sido usado como base para a aplicação de métodos bem como no desenvolvimento de modelos gerais para sistemas dos mais variados tipos.

Neste trabalho, utilizaremos conceitos e métodos de física estatística para investigar sistemas complexos distintos. Especificamente, analisaremos i) a circulação de revistas e jornais[20]; ii) as citações de periódicos científicos[21]; iii) atividades religiosas de uma organização mundial[22]; e iv) um sinal geomagnético[23]. Nossos objetivos são os seguintes: i) caracterizar o comportamento global de cada sistema - sua estrutura hierárquica e sua dinâmica; ii) explorar a presença de invariância de escala e identificar comportamentos universais; iii) desenvolver modelos estocásticos que reproduzam aspectos do comportamento observado.

No capítulo 2, apresentaremos uma breve revisão de conceitos de física estatística que utilizaremos no estudo dos sistemas mencionados acima, incluindo distribuições de probabilidade, correlações temporais e multifractalidade [24-27]. Descreveremos também métodos recentemente desenvolvidos no contexto da física estatística dos sistemas complexos: i) *detrended fluctuation analysis* (DFA) - para a detecção de correlações

de longo alcance[28, 29]; e ii) *multifractal detrended fluctuation analysis* (mf-DFA) - para a detecção de multifractalidade[30].

No capítulo 3, consideraremos exemplos específicos de sistemas complexos que têm sido estudados sob o enfoque da física estatística. No contexto das organizações e atividades humanas, consideraremos o tamanho de firmas, o produto interno bruto (PIB) de países e o tamanho de universidades. Ademais, no contexto da dinâmica de sinais complexos, destacaremos um sinal fisiológico, um sistema hidrodinâmico e o tráfego na Internet.

Utilizando os conceitos e métodos considerados no capítulo 2 e seguindo a linha de pesquisa descrita no capítulo 3, passaremos a investigar, no capítulo 4, a circulação de revistas; no capítulo 5, as citações de periódicos científicos; no capítulo 6, as atividades religiosas e no capítulo 7, o sinal geomagnético. Por fim, as conclusões e as considerações finais serão descritas no capítulo 8, onde destacaremos os exemplos de invariância de escala e universalidade encontrados nos sistemas considerados.

Capítulo 2

Conceitos e métodos

Apresentaremos aqui uma breve revisão de conceitos básicos e métodos utilizados no contexto da física estatística dos sistemas complexos. Especificamente, focalizaremos os conceitos de *i*) densidade de probabilidade; *ii*) correlações de longo alcance; e *iii*) multifractalidade. Descreveremos também um método recentemente desenvolvido para a detecção de correlações de longo alcance em sinais não-estacionários - *detrended fluctuation analysis* (DFA) - e um método para a análise de sinais multifractais não estacionários - *multifractal detrended fluctuation analysis* (mf-DFA).

2.1 Densidade de probabilidade

Os conceitos de variável aleatória e probabilidade são amplamente utilizados no contexto da física estatística dos sistemas complexos[2, 3]. A definição de probabilidade se faz construindo o conjunto de todos os possíveis resultados de uma experiência, agrupando-os em subconjuntos mutuamente excludentes. Se a cada um desses subconjuntos for atribuído um número real não negativo tal que a soma deles seja igual à unidade, então estaremos diante de uma distribuição de probabilidades definida sobre o conjunto dos possíveis resultados[24, 25].

Por exemplo, considere uma variável aleatória discreta l e suponha que a cada valor de l esteja associado um número real $p_l > 0$ tal que $\sum_l p_l = 1$. Caso isso aconteça, p_l será a distribuição de probabilidade associada à variável aleatória l . Considere agora uma variável aleatória contínua x que pode assumir qualquer valor sobre a reta real. Nesse caso, associamos uma probabilidade a cada intervalo da reta. A probabilidade

de que x esteja no intervalo $[a, b]$ pode ser escrita como

$$\int_a^b p(x) dx, \quad (2.1)$$

em que $p(x) > 0$ é a função densidade de probabilidade que deve satisfazer a propriedade

$$\int_{-\infty}^{\infty} p(x) dx = 1. \quad (2.2)$$

No apêndice A consideramos alguns exemplos de distribuições de probabilidade, incluindo as que serão utilizadas neste trabalho.

Os momentos da distribuição $p(x)$ são definidos como

$$\langle x^n \rangle = \int x^n p(x) dx, \quad (2.3)$$

sendo que omitimos os limites de integração, por simplicidade. O valor médio μ e o desvio padrão σ da variável x são definidos como segue:

$$\mu = \langle x \rangle \quad (2.4)$$

e

$$\sigma = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}. \quad (2.5)$$

2.1.1 Histogramas e distribuições acumuladas

Dado um conjunto finito de N realizações da variável x , determinaremos a função densidade de probabilidade $p(x)$ associada à ela. Podemos fazer isso agrupando as realizações da variável x em subconjuntos mutuamente excludentes de tamanho Δx . A cada um destes subconjuntos atribuímos um número real não negativo, $\Delta n/N$, tal que a soma deles seja igual à unidade. A quantidade Δn refere-se ao número de realizações da variável x dentro de um subconjunto Δx . Neste caso, a função densidade de probabilidade pode ser aproximada pela relação

$$p(x) \simeq \frac{\Delta n}{N \Delta x}. \quad (2.6)$$

A função densidade de probabilidade $p(x)$ também pode ser obtida indiretamente

a partir da distribuição acumulada

$$p_c(x) = \int_x^\infty p(y) dy. \quad (2.7)$$

Na prática, o cálculo de $p_c(x)$ envolve obter, para cada valor de x , o número de medidas maiores ou iguais a x .

2.2 Distribuições do tipo lei de potência

Um exemplo digno de nota são as distribuições do tipo lei de potência,

$$p(x) \propto x^{-\alpha}, \quad (2.8)$$

com $\alpha > 0$. Uma propriedade distintiva da lei de potência é que ela satisfaz a equação

$$p(bx) = g(b)p(x), \quad (2.9)$$

para qualquer b constante. Esta relação nos diz que se aumentarmos a escala da variável x por um fator de b , a forma da distribuição não muda, exceto por uma constante multiplicativa. Tal propriedade é chamada de invariância de escala.

Usualmente, no estudo de sistemas complexos, as distribuições do tipo lei de potência aparecem como comportamento assintótico de distribuições mais gerais[2, 3, 8, 9] (veja apêndice A).

2.3 Correlações de longo alcance

O conceito de correlações de longo alcance tem sido largamente utilizado no estudo da dinâmica de sistemas complexos[2, 3]. Dada uma seqüência de variáveis aleatórias $x(i)$, com $i = 1, \dots, N$, é possível obter a função de auto-correlação $A(s)$, definida como

$$A(s) = \langle \bar{x}(i)\bar{x}(i+s) \rangle = \frac{1}{N-s} \sum_{i=1}^{N-s} \bar{x}(i)\bar{x}(i+s), \quad (2.10)$$

com $\bar{x}(i) = x(i) - \langle x \rangle$ e s sendo a escala de tempo.

A função de autocorrelação $A(s)$ é uma medida da correlação entre elementos separados entre si por s passos: $x(i)$ e $x(i + s)$. No caso trivial em que a série não é correlacionada, $A(s) = 0$ para $s > 0$. No caso de séries correlacionadas, $A(s)$ é tipicamente uma função decrescente, partindo de $A(0) = \langle x^2(i) \rangle$ e alcançando $A(s) \simeq 0$ para grandes valores de s .

Se a função $A(s)$ decai rapidamente a zero - por exemplo, exponencialmente - diz-se que a série exibe correlações de curto alcance. Por outro lado, $A(s)$ pode apresentar um decaimento mais lento do tipo lei de potência,

$$A(s) \propto s^{-\gamma}, \quad (2.11)$$

com $0 < \gamma < 1$. Neste caso, diz-se que a série exibe correlações de longo alcance. O cálculo direto da função de auto-correlação $A(s)$ (usando-se a eq. (2.10)) exige que a série seja estacionária. De modo similar, métodos convencionais tais como *Hurst analysis* e *spectral analysis* são indicados para o estudo de séries estacionárias[44, 45].

Estritamente falando, um processo estocástico é estacionário se sua densidade de probabilidade $p(x)$ for invariante sob translação temporal. Em uma definição mais geral (menos estrita), um processo estacionário é definido por três condições: $\langle x(i) \rangle = \mu$, $\langle x(i)x(j) \rangle = A(i, j)$, com $A(i, j) = A(s)$ sendo $s = j - i$ e $\langle x^2(i) \rangle = A(0)$. Isso implica que a variância do processo, $A(0) - \mu^2$, é independente do tempo. Em geral, séries que representam a dinâmica de sistemas complexos não satisfazem tais requisitos. A série pode estar acompanhada de ruídos sobrepostos ao sinal original ou de tendências de origem diversa.

A presença de um comportamento não estacionário pode levar a uma falsa detecção de correlações de longo alcance. Para se obter resultados confiáveis, é preciso distinguir as tendências das flutuações intrínsecas do sistema e então remover as tendências sem alterar o comportamento original das correlações. Métodos tais como *wavelet transform modulus maxima*(WTMM)[46] e *detrended fluctuation analysis*(DFA)[28, 29] foram desenvolvidos justamente com este objetivo. Estes métodos eliminam sistematicamente tendências polinomiais de diferentes ordens, possibilitando medidas confiáveis das correlações. Em particular, o método DFA explora propriedades fractais da série em estudo e é mais simples de ser implementado. A seguir, passaremos a descrever o método DFA, que será utilizado neste trabalho.

2.3.1 O método DFA

O método DFA tem sido amplamente utilizado na detecção de correlações de longo alcance em sistemas complexos, sendo aplicado, por exemplo, no estudo de seqüências de DNA[47], temperatura atmosférica[11, 48], índices econômicos[49], caminhada humana[50], batimentos cardíacos[51], estrutura de nuvens[52], receptores neurais[53], atividade física[54], sinais geoeletricos[55], velocidade do vento[56], sinais geomagnéticos[57], espalhamento por raios x[58] entre muitos outros.

O procedimento envolvido no DFA consiste basicamente de quatro passos[28, 29]: *i*) obter a série acumulada; *ii*) dividir a série acumulada em janelas ou subconjuntos; *iii*) remover a tendência local em cada janela; e *iv*) calcular a função de flutuação em termos do tamanho das janelas.

Para ilustrar de modo mais específico cada um destes passos, considere uma série temporal representada pelos elementos $x(i)$, com $i = 1, \dots, N$. Obtemos a série acumulada cujos elementos são dados por

$$y(i) = \sum_{k=1}^i [x(k) - \langle x \rangle]. \quad (2.12)$$

A subtração da média $\langle x \rangle$ não é obrigatória, visto que ela será eliminada em um dos próximos passos.

Dividimos então a série acumulada em N_s janelas não coincidentes de tamanho s . Quando o número de termos da série não for um múltiplo da escala de tempo s , restará uma parte da série acumulada. Para não desprezar esta parte da série, repetimos o mesmo procedimento a partir do fim da série, obtendo um total de $2N_s$ janelas de tamanho s . Dentro de cada janela, indexada pelo índice ν , ajustamos um polinômio de ordem l aos dados, representando a tendência local. Costuma-se usar os termos DFA-1, se $l = 1$, DFA-2, se $l = 2$ e assim por diante. A seguir, subtraímos a série acumulada de sua tendência local, calculando as diferenças

$$w_s(i) = y(i) - y_\nu(i), \quad (2.13)$$

em que y_ν é o polinômio ajustado aos dados dentro do segmento ν . A figura (2.1) ilustra este procedimento partindo-se da série original até chegar na série das diferenças.

Como as diferenças $w_s(i)$ são obtidas por uma subtração da série acumulada de seus respectivos polinômios de ajuste, a ordem do polinômio, obviamente, relaciona-se com

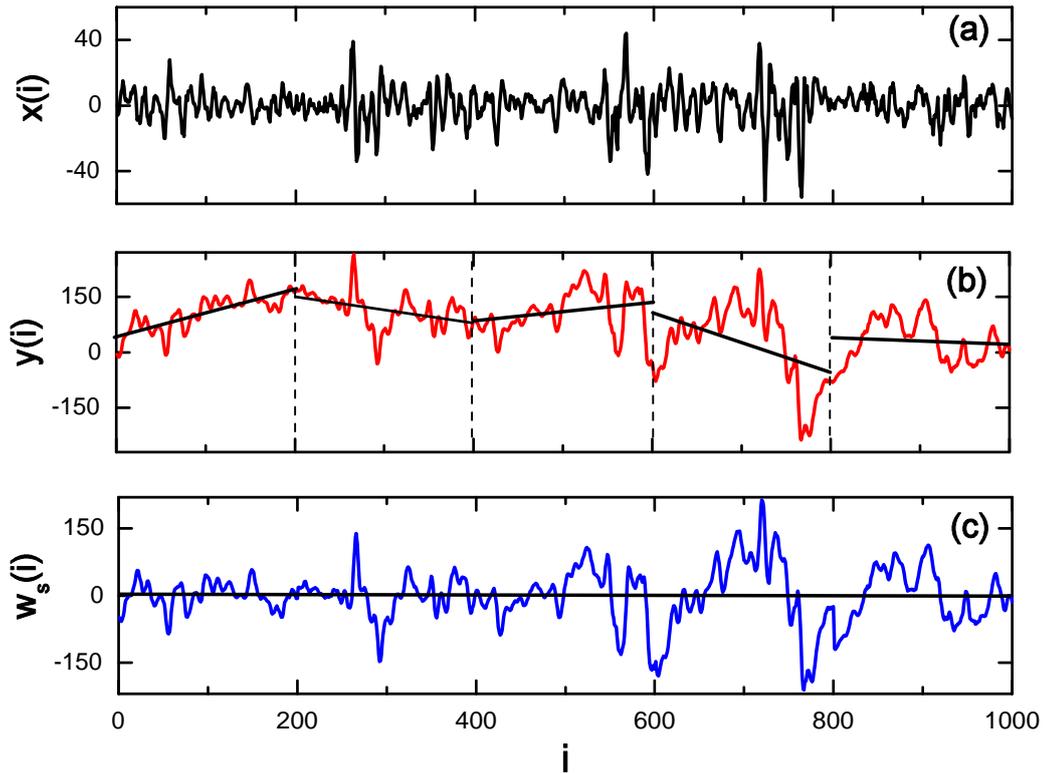


Figura 2.1: Ilustração de procedimentos envolvidos no método DFA. **(a)** Série temporal $x(i)$ com $N = 1000$ termos. **(b)** Série acumulada $y(i)$ obtida a partir da série original, cujos elementos são dados pela eq. (2.12). As linhas tracejadas na vertical indicam janelas de tamanho $s = 200$ e as linhas retas dentro de cada janela são polinômios de ordem 1 ajustados aos dados (por mínimos quadrados) representando a tendência local. **(c)** Série das diferenças $w_s(i)$ entre a série acumulada e as tendências locais, cujos elementos são dados pela eq. (2.13), com $s = 200$.

sua capacidade de eliminar tendências. De fato, DFA- l elimina tendências de ordem l na série somada e de ordem $l - 1$ na série original.

O próximo passo consiste em se calcular a variância,

$$F_s^2(\nu) = \frac{1}{s} \sum_{i=1}^s w_s^2[(\nu - 1)s + i], \quad (2.14)$$

fazendo uma média de $w_s^2(i)$ sobre todos os dados dentro do segmento ν . Finalmente, definimos a função de flutuação

$$\tilde{F}(s) = \left[\langle F_s^2(\nu) \rangle \right]^{1/2}, \quad (2.15)$$

onde $\langle F_s^2(\nu) \rangle$ é a média de $F_s^2(\nu)$ calculada sobre todos os $2N_s$ segmentos, definida como

$$\langle F_s^2(\nu) \rangle = \frac{1}{2N_s} \sum_{\nu=1}^{2N_s} F_s^2(\nu). \quad (2.16)$$

Em geral, $\tilde{F}(s)$ será uma função crescente visto que a variância tipicamente aumenta com o tamanho das janelas. Um comportamento do tipo lei de potência,

$$\tilde{F}(s) \sim s^h, \quad (2.17)$$

indica que a série acumulada apresenta fractalidade ou auto-similaridade.

2.3.2 Detectando correlações com DFA

Conforme já comentado, quando a série considerada exibe fractalidade, $\tilde{F}(s) \sim s^h$. O expoente h pode ser obtido simplesmente por um ajuste linear de $\tilde{F}(s)$ por s na escala logarítmica. Dada uma série com N termos, qual o intervalo, $s_{\min} < s < s_{\max}$, mais apropriado para obter o expoente h ? Como s refere-se ao tamanho das janelas em que a série acumulada é dividida, podemos considerar dois extremos: *i*) se s_{\max} for muito grande, aproximando-se do comprimento da série, teremos um número pequeno de janelas; e *ii*) se s_{\min} é muito pequeno teremos poucos dados em cada janela. Como ambos os extremos são indesejáveis, geralmente usa-se $s_{\max} \sim N/10$ (um décimo do comprimento da série) e $s_{\min} \geq 6$ [59].

Entretanto, o cálculo de h para séries que exibem $h < 0,5$ exige que s_{\min} assumam valores muito maiores que ≥ 6 . Pode-se contornar esta dificuldade por se aplicar DFA,

não sobre a série original, mas sobre a série acumulada[59] (neste caso a série, por fim, acaba sendo integrada duas vezes). Neste caso, a nova função de flutuação $F(s)$ será proporcional a $s\tilde{F}(s)$. Sendo assim, se $\tilde{F}(s)$ comporta-se como lei de potência com expoente h (veja eq. (2.17),

$$F(s) \sim s^{h+1}. \quad (2.18)$$

Desse modo, espera-se que a série caracterizada por um certo expoente $h_0 < 0,5$ passe a exibir um expoente maior que $0,5$ ($h = h_0 + 1$), eliminando a dificuldade anteriormente mencionada relacionada com o valor de s_{min} .

O expoente h , também conhecido como expoente de Hurst, pode revelar informações relevantes sobre as correlações na série em estudo. Consideremos os seguintes casos:

- Se a série em estudo não for correlacionada (ruído branco), espera-se que $h = 0,5$ [60].
- Se a série exibe correlações de curto alcance, do tipo exponencial ou mesmo do tipo lei de potência, mas com expoente $\gamma > 1$, h poderá ser diferente de $0,5$ para pequenas escalas de tempo s e se aproximará de $0,5$ para s grande[60].
- Se a série apresenta correlações de longo alcance, do tipo lei de potência com expoente $0 < \gamma < 1$, espera-se que $0,5 < h < 1$ [60]. Neste caso, o expoente h pode ser relacionado com o expoente γ pela relação $h = 1 - \gamma/2$, e a série é persistente. Por outro lado, valores de h no intervalo $0 < h < 0,5$ indica a presença de anti-correlações de longo alcance e a série é anti-persistente;
- O caso $h = 1$ corresponde ao chamado ruído $1/f$. Se $h > 1$ existem correlações mas não na forma de lei de potência. O caso $h = 1,5$ indica ruído browniano, a integração do ruído branco ($h = 0,5$). O expoente h pode ser visto como um indicador da “rugosidade” da série original. Quanto maior o valor de h , mais “lisa” será a série. Deste ponto de vista, $h = 1$ pode ser interpretado como um estado intermediário entre o aspecto muito rugoso do ruído branco e o aspecto mais liso do ruído browniano.

Costuma-se repetir a análise com DFA para versões embaralhadas da série original. No procedimento de embaralhar a série, os seus valores são colocados em ordem aleatória, destruindo possíveis correlações. Por este motivo, espera-se obter expoentes $h \simeq 0,5$. Na fig. (2.2) ilustramos a aplicação do DFA para uma série anti-persistente

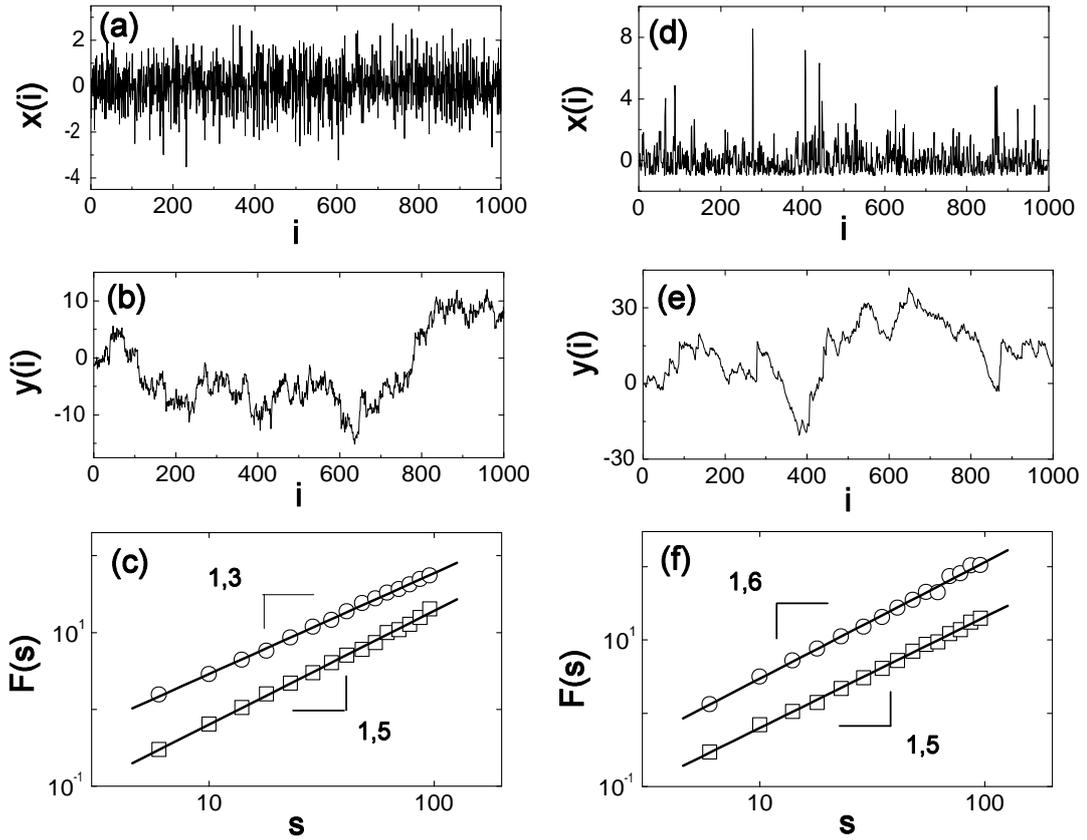


Figura 2.2: Detectando correlações com DFA. **(a)** Exemplo de série anti-persistente. **(b)** Série em (a) integrada. **(c)** Função de flutuação $F(s)$ (círculos) obtida pela aplicação do método DFA de segunda ordem na série em (b). O ajuste linear fornece inclinação $\simeq 1,3$, indicando que $h \simeq 0,3$ (compare com eq. (2.18)). O mesmo procedimento aplicado a uma versão embaralhada da série em (a) fornece $h \simeq 0,5$ (quadrados). **(d)** Exemplo de série persistente. **(e)** Série em (d) integrada. **(f)** Função de flutuação $F(s)$ (círculos). Neste caso, a inclinação é $\simeq 1,6$, indicando que $h \simeq 0,6$. Novamente, $h \simeq 0,5$ para a versão embaralhada (quadrados).

($h < 0,5$) e para uma série persistente ($h > 0,5$). Observe que, em ambos os casos, os expoentes assumem valores $h \simeq 0,5$ para suas versões embaralhadas.

O método DFA, naturalmente, tem suas limitações. Algumas séries talvez tenham tendências que o método não consegue remover. Nesta direção, séries geradas artificialmente que apresentam diversos tipos de comportamento não estacionário têm sido sistematicamente investigados com o DFA[59, 61]. Estes estudos incluem séries com diferentes tipos de tendência[59] (polinomiais, senoidais, leis de potência), séries com segmentos removidos (descontinuidades), com saltos aleatórios e com diferentes comportamentos locais (desvio padrão, correlações)[61]. Tais estudos podem ser úteis na correta interpretação dos resultados obtidos da aplicação do DFA.

A seguir, vamos considerar aspectos do conceito de multifractalidade em séries temporais. Descreveremos também uma generalização do método DFA para a análise de sinais multifractais não estacionários.

2.4 Multifractalidade

O conceito de multifractalidade também têm sido amplamente explorado no estudo dos sistemas complexos[26, 27]. Para uma definição de série multifractal, voltemos ao contexto das correlações temporais de longo alcance. Na seção anterior, vimos uma maneira de calcular o expoente γ da relação $A(s) = \langle x(i)x(j) \rangle \sim s^{-\gamma}$, com $j = i + s$, para uma dada série. No entanto, uma descrição mais completa do processo estocástico talvez envolva o cálculo de funções de correlação de ordem m , com $m > 2$, tais como $\langle x(i)x(j)x(k) \rangle$ ou $\langle x(i)x(j)x(k)x(l) \rangle$ e seus respectivos expoentes γ_m .

Séries em que os expoentes γ_m são linearmente dependentes de γ_2 são chamadas “lineares” ou “monofractais”, visto que um único expoente é necessário para caracterizá-las. Por outro lado, séries “não-lineares” ou “multifractais” são aquelas em que os expoentes γ_m não são linearmente dependentes de γ_2 , tornando necessário o cálculo de uma família ou espectro de expoentes[26, 27].

A análise multifractal de uma série estacionária com elementos $x(i)$ envolve o cálculo da função de partição $Q_m(s) = \langle |x(i) - x(i + s)|^m \rangle$, em que o valor médio é calculado sobre toda a série. Nos casos em que $Q_m(s) \sim s^{\tau(m)}$, a forma da dependência de $\tau(m)$ com m caracteriza a multifractalidade (ou monofractalidade) da série. Séries monofractais são caracterizadas por um comportamento linear de $\tau(m)$ com m , enquanto que um comportamento não linear de $\tau(m)$ com m indica a presença

de multifractalidade[26, 27].

Este tipo de análise multifractal é válido apenas para séries estacionárias. Por outro lado, dois métodos se destacam na detecção de multifractalidade em séries não estacionárias: *wavelet transform modulus maxima*[46](WTMM) e *multifractal detrended fluctuation analysis*[30](mf-DFA). De modo geral, ambos os métodos fornecem resultados equivalentes[30, 62, 63]. A seguir, passamos a descrever o método mf-DFA, visto que ele é uma generalização do já considerado método DFA.

2.4.1 O método mf-DFA

O método *multifractal detrended fluctuation analysis*[30](mf-DFA) é uma generalização do já comentado *detrended fluctuation analysis*[28, 29](DFA). Em anos recentes, o método mf-DFA tem sido aplicado no estudo de sinais multifractais em diversos campos, tais como velocidade do vento[56], vazão de rios[62], índices econômicos [64-66], sinais geoeletricos[67], sinais geomagnéticos[68, 69], terremotos[70], pressão sanguínea[71], camada de ozônio[72], turbulência[73], manchas solares[74] entre outros.

O procedimento envolvido no mf-DFA consiste basicamente de quatro passos[30]: *i*) obter a série acumulada; *ii*) dividir a série acumulada em janelas ou subconjuntos; *iii*) remover a tendência local em cada janela; e *iv*) calcular a função de flutuação generalizada em termos do tamanho das janelas.

Para ilustrar de modo mais específico cada um destes passos, considere uma série temporal representada pelos elementos $x(i)$, com $i = 1, \dots, N$. Os passos *i*, *ii* e *iii* são idênticos aos três primeiros passos descritos na seção 2.2.1 para o método DFA.

A diferença entre os métodos aparece no próximo passo, relacionado ao cálculo da função de flutuação. A partir da série das diferenças cujos elementos são dados pela eq. (2.13), calcula-se a variância $F_s^2(\nu)$ definida na eq. (2.14). A função de flutuação generalizada é definida como

$$\tilde{F}_m(s) = [\langle F_s^2(\nu) \rangle_m]^{1/m}, \quad (2.19)$$

com

$$\langle F_s^2(\nu) \rangle_m = \frac{1}{2N_s} \sum_{\nu=1}^{2N_s} [F_s^2(\nu)]^{m/2}, \quad (2.20)$$

em analogia com as equações (2.15) e (2.16). Observe que se $m = 2$ o procedimento recai no método DFA padrão.

Se a série exibe propriedades fractais (ou multifractais), $\tilde{F}_m(s)$ crescerá como uma lei de potência

$$\tilde{F}_m(s) \sim s^{h(m)}, \quad (2.21)$$

sendo $h(m)$ chamado de expoente de Hurst generalizado ($h(2) = h$). O valor $h(0)$ corresponde ao limite de $h(m)$ para $m \rightarrow 0$, e não pode ser determinado diretamente usando eq. (2.21). Ao invés disso, $h(0)$ é obtido empregando-se a aproximação[30]

$$\tilde{F}_0(s) = \exp \left[\frac{1}{4N_s} \sum_{\nu=1}^{2N_s} \ln[F_s^2(\nu)] \right] \sim s^{h(0)}. \quad (2.22)$$

Assim como no caso do método DFA, costuma-se aplicar mf-DFA na série acumulada ao invés da original para evitar problemas com sinais anti-persistentes[30]. Assim sendo, a função de flutuação generalizada obtida a partir da série integrada será dada por,

$$F_m(s) \sim s^{h(m)+1}, \quad (2.23)$$

em analogia com a eq. 2.18.

2.4.2 Detectando multifractalidade com mf-DFA

Nas séries monofractais, o comportamento das variâncias $F_s^2(\nu)$, dentro de uma janela ν , é idêntico para pequenas e grandes janelas. Assim, quando se toma a média dada pela eq. (2.20), o resultado será praticamente independente de m . Por outro lado, se pequenas e grandes flutuações exibirem diferentes propriedades, então isso se refletirá numa dependência de $h(m)$ com m . Por exemplo, se $m > 0$, segmentos com grandes flutuações dominarão a média $F_m(s)$ e $h(m)$ descreverá o comportamento dos segmentos com grande variância. Por outro lado, se $m < 0$, a média $F_m(s)$ será dominada pelos segmentos caracterizados por pequenas flutuações e $h(m)$ descreverá o comportamento dos segmentos com pequena variância. Usualmente, grandes flutuações são caracterizadas por pequenos expoentes e pequenas flutuações por grandes expoentes. Isso faz com que $h(m)$ tenha uma dependência significativa com m . Portanto, séries monofractais são caracterizadas por $h(m) \sim cte$. Por outro lado, uma dependência significativa de $h(m)$ com m indica a presença de multifractalidade[30].

O expoente generalizado $h(m)$ relaciona-se ao expoente (ou espectro) multifractal

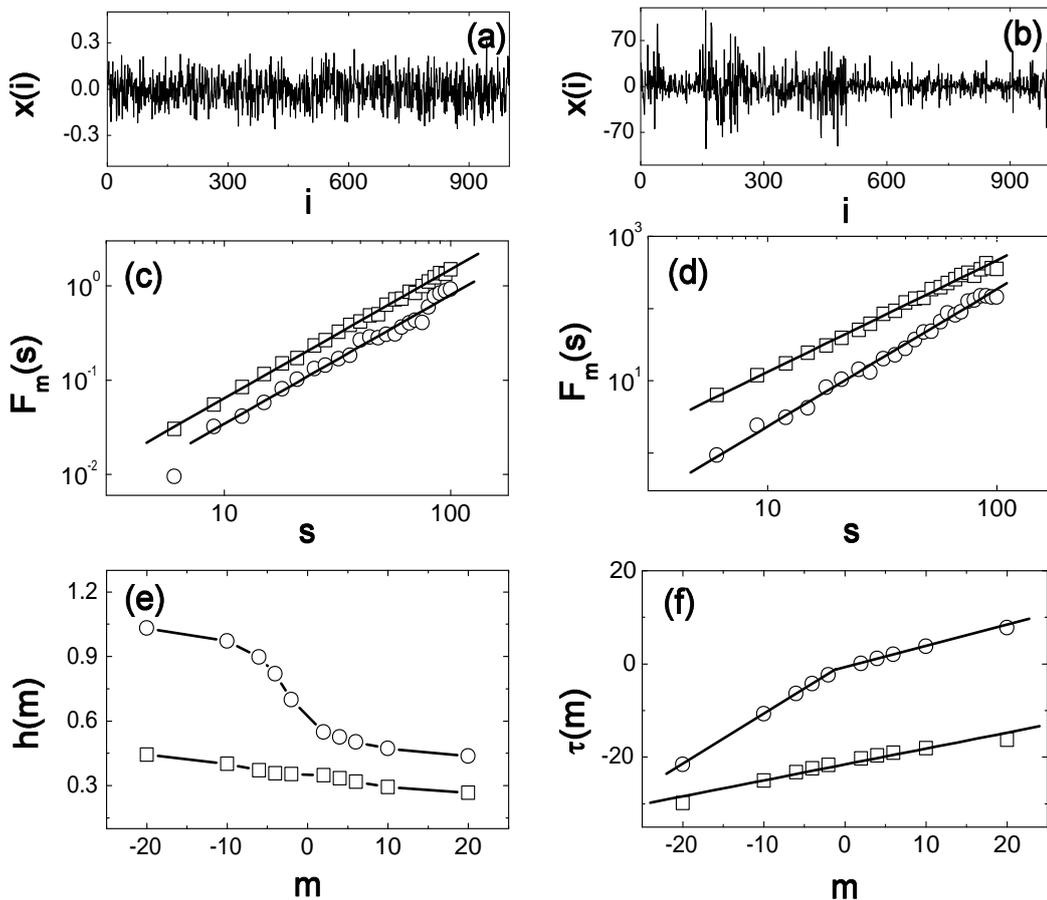


Figura 2.3: Detectando multifractalidade com mf-DFA (de segunda ordem). **(a)** Exemplo de série monofractal. **(b)** Exemplo de série multifractal. **(c)** Função de flutuação $F_m(s)$, com $m = -6$ (círculos) e $m = 2$ (quadrados), para a série monofractal em (a). **(d)** Função de flutuação $F_m(s)$, com $m = -6$ (círculos) e $m = 2$ (quadrados), para a série multifractal em (b). **(e)** Expoentes $h(m)$ versus m para a série monofractal (quadrados) e para a série multifractal (círculos). **(f)** Exponentes $\tau(m)$ versus m para a série monofractal (quadrados) e para a série multifractal (círculos).

$\tau(m)$ pela relação[30]

$$\tau(m) = m h(m) - 1. \quad (2.24)$$

Na prática, a detecção de multifractalidade é melhor visualizada pela dependência de $\tau(m)$ com m . Séries monofractais são caracterizadas por um comportamento linear de $\tau(m)$ com m , enquanto que um comportamento não linear de $\tau(m)$ com m indica a presença de multifractalidade[30].

Para ilustrar a detecção de multifractalidade, considere as séries ilustradas nas figs. 2.3a e 2.3b. Observe que a forte dependência do expoente $h(m)$ com m e o comportamento não linear de $\tau(m)$ versus m , para a série da fig. 2.3b, indicam a presença de multifractalidade.

Dois diferentes tipos de multifractalidade podem estar presentes numa série: *i*) devido ao tipo de distribuição dos valores da série; ou *ii*) devido a diferentes correlações de longo alcance presentes na série. No que se refere a distinguir entre estes dois tipos de multifractalidade, é útil analisar, além da série original, sua correspondente série embaralhada. No processo de se embaralhar a série, seus valores são colocados em ordem aleatória, e portanto, suas correlações são destruídas. Deste modo, no caso de multifractalidade do tipo *ii*, as séries embaralhadas não apresentarão comportamento multifractal $h(m) \simeq 0,5$. Por outro lado, para multifractais do tipo *i*, não observaremos mudanças visto que a distribuição de probabilidades dos valores da série não é afetada pelo processo de embaralhamento. Se ambos os tipos de multifractalidade estão presentes numa dada série, a série embaralhada apresentará multifractalidade mais fraca do que a série original.

No que segue, ilustraremos como os conceitos e métodos apresentados neste capítulo têm sido aplicados no estudo da dinâmica de sistemas complexos.

Capítulo 3

Física estatística dos sistemas complexos: recentes aplicações

Conceitos e métodos de física estatística têm sido aplicados no estudo de uma grande variedade de sistemas complexos, incluindo organizações e atividades humanas e sinais fisiológicos. Em particular, descreveremos aqui resultados recentemente obtidos no estudo da dinâmica *i)* de atividades econômicas e de pesquisa científica; e *ii)* de um sinal fisiológico relacionado ao coração humano. Destacaremos exemplos notáveis de invariância de escala e comportamento universal nestes temas.

3.1 Dinâmica de crescimento de organizações

Um problema clássico no contexto de organizações econômicas é o estudo do tamanho de firmas [75-78]. Em 1931, o economista francês Gibrat propôs um modelo simples para descrever a dinâmica de crescimento de firmas [75]. O modelo de Gibrat prevê que *i)* o tamanho das firmas segue uma distribuição log-normal; *ii)* as taxas de crescimento seguem uma distribuição Gaussiana; e *iii)* a média e as flutuações das taxas de crescimento são independentes do tamanho da firma.

Em contraste com tais previsões do modelo de Gibrat, evidências empíricas indicam que as taxas de crescimento de firmas apresentam distribuição não-Gaussiana e que o desvio padrão das taxas de crescimento depende do tamanho da firma. A seguir, apresentaremos recentes aplicações da física estatística dos sistemas complexos no estudo

da dinâmica de atividades econômicas, incluindo tamanho de firmas e atividades de pesquisa científica.

3.1.1 Atividades econômicas

Considere os resultados descritos nas refs. [79, 80] e ilustrados na fig. 3.1. Tais resultados referem-se ao conjunto das firmas americanas (manufaturas) entre os anos 1975 a 1991, sendo que o total das vendas (em dólares) define o tamanho da firma. A taxa anual de crescimento é definida como[79]

$$R(t) = \log \left[\frac{S(t+1)}{S(t)} \right], \quad (3.1)$$

com $S(t)$ e $S(t+1)$ sendo o tamanho da firma nos anos t e $t+1$, respectivamente.

A distribuição dos tamanhos, $p(S)$, é aproximadamente log-normal[80]. Ademais, a densidade de probabilidade condicional das taxas anuais de crescimento, $p(R|S)$, apresenta comportamento não Gaussiano (figs. 3.1a e 3.1b), sendo bem ajustada pela distribuição (exponencial)

$$p(R|S) = \frac{1}{\sqrt{2}\sigma(S)} \exp \left(-\frac{\sqrt{2}|R - \mu|}{\sigma(S)} \right), \quad (3.2)$$

com μ e σ sendo a média e o desvio padrão de R . Observe que a distribuição das taxas de crescimento exibe a mesma forma funcional para firmas de diferentes tamanhos - um tipo de invariância de escala[79]. Ademais, o desvio padrão das taxas de crescimento decresce com o tamanho das firmas como uma lei de potência (fig. 3.1c),

$$\sigma(S) \propto S^{-\beta}, \quad (3.3)$$

com $\beta \simeq 0,15$.

As eqs. 3.2 e 3.3 descrevem satisfatoriamente a dinâmica de quatro medidas adicionais de tamanho de firmas (veja figs. 3.1c e 3.1d): *i*) funcionários ($\beta \simeq 0,16$); *ii*) mercadorias ($\beta \simeq 0,16$); *iii*) ativo ($\beta \simeq 0,17$); e *iv*) propriedades e equipamentos ($\beta \simeq 0,18$)[79].

Algo notável nas eqs. 3.2 e 3.3 é que elas governam as taxas de crescimento de um conjunto diverso de firmas, as quais diferem não apenas em tamanho, mas também no que produzem. Isso sugere que os processos que governam a dinâmica de crescimento

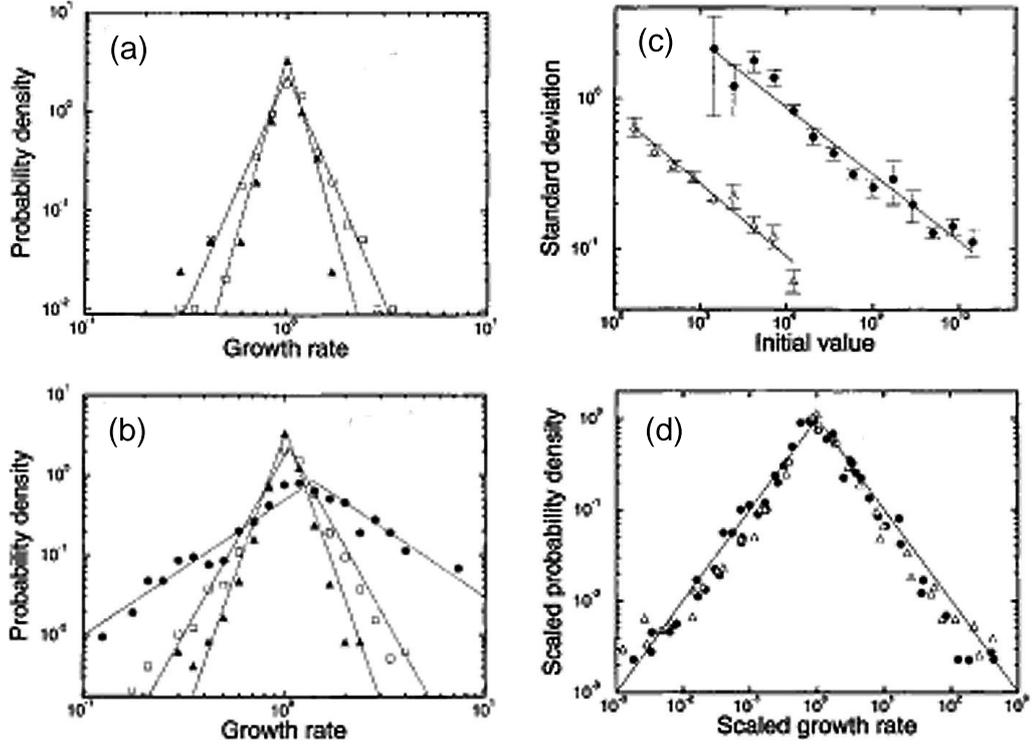


Figura 3.1: Dinâmica de crescimento de firmas. Fonte: ref. [79]. (a) Densidade de probabilidade (*Probability density*) da taxa de crescimento (*Growth rate*), referente aos anos 1990 e 1991, para firmas de diferentes tamanhos (vendas): $4^{11,5} < S < 4^{12,5}$ (quadrados) e $4^{14,5} < S < 4^{15,5}$ (triângulos). (b) Densidade de probabilidade (*Probability density*), referentes ao período 1975-1991, para três grupos de firmas: $4^{8,5} < S < 4^{9,5}$ (círculos), $4^{11,5} < S < 4^{12,5}$ (quadrados) e $4^{14,5} < S < 4^{15,5}$ (triângulos). As linhas sólidas em (a) e (b) representam a distribuição definida na eq. (3.2). (c) Desvio padrão (*Standard deviation*) das taxas de crescimento em termos do tamanho inicial das firmas (*Initial value*), calculado para duas medidas de tamanho: vendas (círculos) e número de funcionários (triângulos). As linhas sólidas representam leis de potência com expoente $\beta \simeq 0,15$ para vendas e $\beta \simeq 0,16$ para funcionários. (d) Densidade de probabilidade (*Scaled probability density*) da taxa de crescimento normalizada (*Scaled growth rate*), para vendas (círculos) e número de funcionários (triângulos).

de produtoras de automóveis, por exemplo, podem ser os mesmos que governam firmas farmacêuticas. Este resultado faz lembrar o conceito de universalidade em física estatística, em que diferentes sistemas podem ser caracterizados pelas mesmas leis globais, independentemente das particularidades do sistema[79].

Embora a distribuição do tamanho de firmas seja aproximadamente log-normal[79], há evidências de que as grandes firmas fogem deste comportamento. De fato, a distribuição de tamanho das grandes firmas exhibe um comportamento do tipo lei de potência, com $\alpha \simeq 2,0$ [81]. Ademais, o desvio padrão das taxas de crescimento de grandes firmas não apresenta uma dependência significativa com o tamanho da firma, o que equivale dizer que $\beta \simeq 0$ [81].

Estudos adicionais indicam que a dinâmica de crescimento do produto interno bruto (PIB) de países segue leis quantitativamente similares às descritas pelas eqs. 3.2 e 3.3[82]. Esta similaridade inclui o valor do expoente β , sendo $\beta \simeq 0,15$ para firmas e países[82].

3.1.2 Atividades de pesquisa científica

O estudo da dinâmica de organizações econômicas, exemplificado na seção anterior, tem sido estendido para o estudo de organizações relacionadas a atividades de pesquisa científica, tais como tamanho de universidades e institutos. Considere, por exemplo, resultados descritos na ref. [83] e ilustrados na fig. 3.2. Os dados analisados referem-se ao fundo para pesquisas *R&D* (*Research and Development*) de 719 universidades americanas no período 1979-95.

Os principais resultados indicam que *i*) o fundo *R&D* das universidades americanas segue uma distribuição log-normal (fig. 3.2a); *ii*) a distribuição das taxas de crescimento é aproximadamente exponencial, cuja forma funcional se mantém para diferentes valores do *R&D* (fig. 3.2b e 3.2d); e *iii*) o desvio padrão das taxas de crescimento decresce com o *R&D* seguindo uma lei de potência, com expoente $\beta \simeq 0,25$ (fig. 3.2c).

Além do fundo para pesquisas *R&D*, foram analisadas medidas adicionais de tamanho das universidades americanas: número de artigos publicados e número de patentes. Ademais, a análise foi repetida para universidades inglesas e canadenses. O análogo da fig. 3.2 também vale para os demais conjuntos de dados considerados, incluindo o expoente $\beta \simeq 0,25$ [83].

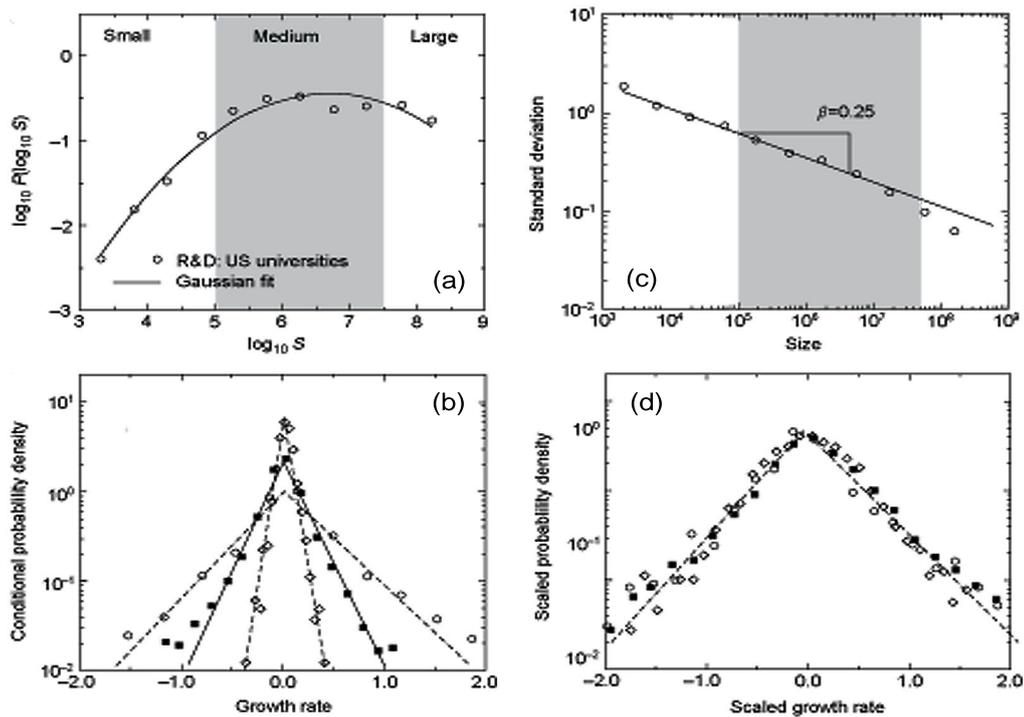


Figura 3.2: Dinâmica de crescimento de universidades. Fonte: ref. [83]. (a) Distribuição $p(\log S)$, sendo S o valor do fundo $R\&D$ em dólares para 719 universidades americanas no período 1979-1995. A linha sólida representa uma distribuição Gaussiana indicando que $p(S)$ é aproximadamente log-normal. (b) Densidade de probabilidade condicional (*Conditional probability density*) das taxas de crescimento anuais (*Growth rate*) para os três grupos de universidades definidos em (a). (c) Desvio padrão (*Standard deviation*) das taxas de crescimento em função do tamanho (*Size*). A linha sólida representa uma lei de potência com expoente $\beta = 0,25$. (d) Densidade de probabilidade (*Scaled probability density*) em termos das taxas de crescimento normalizadas (*Scaled growth rate*) para os três grupos de universidades definidos em (a). As três distribuições são descritas por uma só curva.

Estudos adicionais indicam a presença de padrões análogos na dinâmica de crescimento de outras atividades de pesquisa científica, tais como número de artigos em institutos europeus e em universidades chinesas [84-86].

3.1.3 Características universais

Os resultados descritos nas duas seções anteriores sugerem surpreendentes similaridades entre atividades econômicas e atividades de pesquisa científica[83] (compare as figs. 3.1 e 3.2). Os principais resultados descritos nas seções anteriores são resumidos nas tabs. 3.1 e 3.2.

Tabela 3.1 - Dinâmica de crescimento de atividades econômicas: tamanho de firmas e produto interno bruto (PIB) de países.

	X	$p(\log X)$	$\sigma(X)$	β	$p(R X)$
Firmas	vendas	Gaussiana	$\propto X^{-\beta}$	0,15	exponencial
Firmas	funcionários	Gaussiana	$\propto X^{-\beta}$	0,16	exponencial
Firmas	mercadorias	Gaussiana	$\propto X^{-\beta}$	0,16	exponencial
Firmas	ativo	Gaussiana	$\propto X^{-\beta}$	0,17	exponencial
Firmas	propriedades	Gaussiana	$\propto X^{-\beta}$	0,18	exponencial
Países	PIB	Gaussiana	$\propto X^{-\beta}$	0,15	exponencial

Tabela 3.2 - Dinâmica de crescimento de atividades de pesquisa científica: universidades americanas, inglesas e canadenses.

	X	$p(\log X)$	$\sigma(X)$	β	$p(R X)$
EUA	fundo $R\&D$	Gaussiana	$\propto X^{-\beta}$	0,25	exponencial
EUA	artigos	Gaussiana	$\propto X^{-\beta}$	0,25	exponencial
EUA	patentes	Gaussiana	$\propto X^{-\beta}$	0,25	exponencial
Inglaterra	fundo de pesquisas	Gaussiana	$\propto X^{-\beta}$	0,25	exponencial
Canadá	fundo de pesquisas	Gaussiana	$\propto X^{-\beta}$	0,25	exponencial

Esses resultados sugerem que atividades econômicas e atividades de pesquisa científica exibem o mesmo padrão de crescimento. Especificamente, em ambos os casos, *i*) a distribuição dos tamanhos é aproximadamente log-normal; *ii*) a distribuição das taxas de crescimento exibe decaimento exponencial, independentemente do tamanho; e *iii*) o desvio padrão das taxas de crescimento decresce com o tamanho seguindo uma lei de potência. Entretanto, note que os expoentes β assumem valores distintos para atividades econômicas e atividades de pesquisa.

Estas similaridades sugerem que organizações distintas - tais como firmas, países e universidades - poderiam ser governadas por mecanismos de crescimento em comum[82, 83]. Esta possibilidade é consistente com o conceito de universalidade em sistemas complexos, que sugere a existência de mecanismos “universais” que produzem leis gerais independentes dos detalhes do sistema[82, 83].

Ademais, o intrigante fato de que universidades se comportam como firmas sugere interessantes analogias entre atividades econômicas e atividades de pesquisa científica. Por exemplo, a revisão dos artigos por especialistas e o direcionamento do governo referentes à pesquisa científica poderia levar a resultados similares àqueles induzidos por forças de mercado. O análogo da revisão por especialistas seria o controle de qualidade das mercadorias e o análogo do direcionamento do governo seria a regulamentação ou controle dos produtos[83].

A seguir, citaremos exemplos de modelos estocásticos desenvolvidos para descrever a dinâmica de organizações.

3.1.4 Modelos estocásticos

O modelo de Gibrat para o tamanho de firmas, já mencionado, baseia-se nas seguintes hipóteses[75]: *i*) a taxa de crescimento de uma firma é independente de seu tamanho; *ii*) as taxas de crescimento sucessivas de uma firma não apresentam correlações temporais; e *iii*) as firmas não interagem. Na forma matemática, este modelo pode ser expresso pelo processo estocástico multiplicativo[80]

$$S(t + 1) - S(t) = \varepsilon(t)S(t), \quad (3.4)$$

sendo $S(t + 1)$ e $S(t)$ os tamanhos da firma nos tempos $t + 1$ e t , e $\varepsilon(t)$ um número aleatório não correlacionado seguindo uma distribuição de probabilidades Gaussiana com variância muito menor que a unidade.

Visto que o modelo de Gibrat falha em reproduzir os resultados empíricos descritos pelas eqs. (3.2) e (3.3), poderíamos nos perguntar se modificações neste modelo simples poderiam descrever mais precisamente o comportamento observado. Nesta direção, uma alteração do modelo de Gibrat foi apresentada na ref. [79], gerando distribuições consistentes com a eq. (3.2). Entretanto, o modelo apresentado falha em reproduzir o decaimento do tipo lei de potência descrito pela eq. (3.3)[80].

O modelo de Gibrat considera cada organização como uma unidade sem estrutura interna[75]. Entretanto, organizações tais como firmas, países e universidades costumam apresentar estrutura interna complexa. Por exemplo, firmas geralmente são compostas de divisões, de modo que o tamanho de uma firma costuma ser dado pela soma dos tamanhos das suas subunidades. Modelos tais como os propostos nas refs. [87, 88] levam em conta a estrutura interna de cada organização e, deste modo, reproduzem o comportamento descrito pelas eqs. (3.2) e (3.3). Outros tipos de modelo têm sido apenas parcialmente bem sucedidos em reproduzir o comportamento observado[89].

3.2 Dinâmica dos sinais complexos

O estudo da dinâmica de sinais complexos dos mais variados tipos, sob o enfoque da física estatística, envolve o cálculo de propriedades tais como *i*) distribuição das flutuações; *ii*) correlações de longo alcance; e *iii*) multifractalidade. Em particular, conceitos e métodos de física estatística têm sido adaptados e estendidos para o estudo de sinais fisiológicos. Nesse contexto, destaca-se o estudo dos intervalos de tempo entre batimentos cardíacos. No que segue, consideraremos recentes resultados obtidos neste tema.

3.2.1 Sinal fisiológico

Os intervalos de tempo entre batimentos cardíacos de indivíduos saudáveis pertencem a uma classe de sinais complexos que exibem correlações de longo alcance [90-92], multifractalidade[93, 94] e distribuições não-Gaussianas[91, 95, 96].

Considere os resultados descritos na ref. [91] e ilustrados na fig. 3.3. Na fig. 3.3, vemos a série dos intervalos de tempo entre batimentos cardíacos de um indivíduo saudável (fig. 3.3a) e de um indivíduo doente (fig. 3.3b). As flutuações no sinal

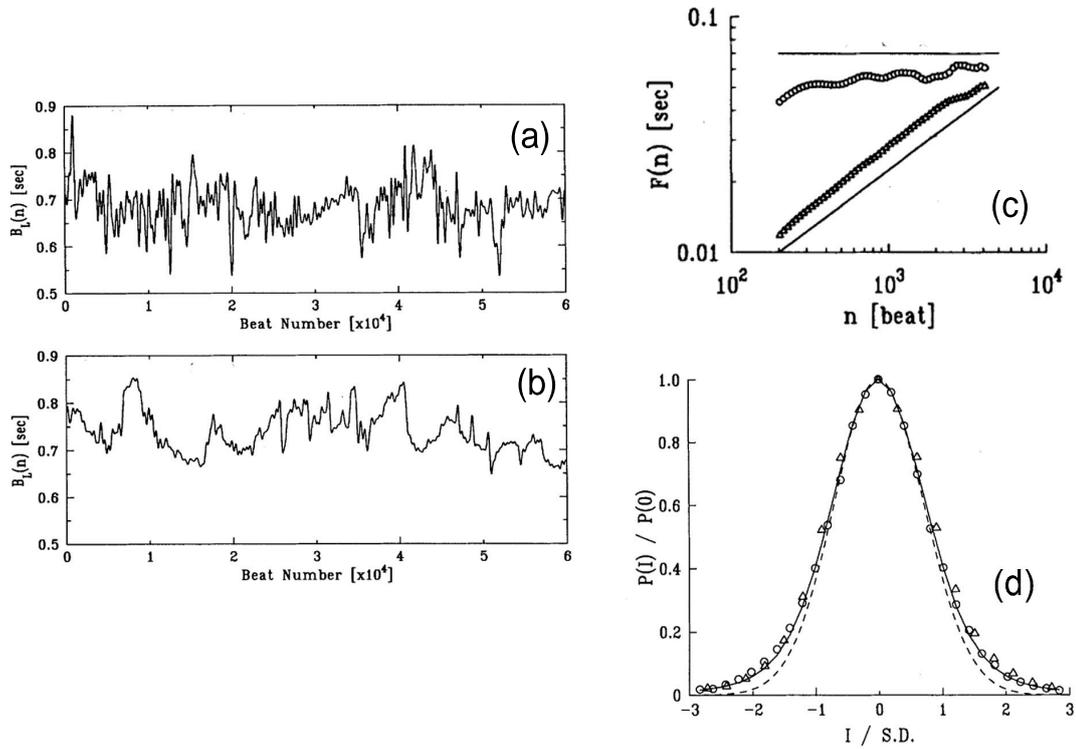


Figura 3.3: Correlações de longo alcance e distribuição não-Gaussiana na dinâmica dos batimentos cardíacos. Fonte: ref. [91]. Intervalos de tempo entre as batidas do coração, $B_L(n)$, para (a) um indivíduo saudável e (b) um indivíduo com doença cardíaca. (c) Função de flutuação $F(n)$, definida na ref. [91], para as séries em (a) e (b). Observe que $F(n) \sim n^\delta$, com $\delta \simeq 0$ para o indivíduo saudável e $\delta \simeq 0,5$ para o indivíduo doente. (d) Distribuição dos incrementos sucessivos, $p(I)/p(0)$, com $I(n) = B_L(n + 1) - B_L(n)$, para os indivíduos saudável (círculos) e doente (triângulos). Para facilitar a comparação, os incrementos são divididos pelo seu desvio padrão. A linha sólida representa uma distribuição de Lévy com $\psi = 1,7$ e a linha tracejada corresponde a uma distribuição Gaussiana ($\psi = 2,0$).

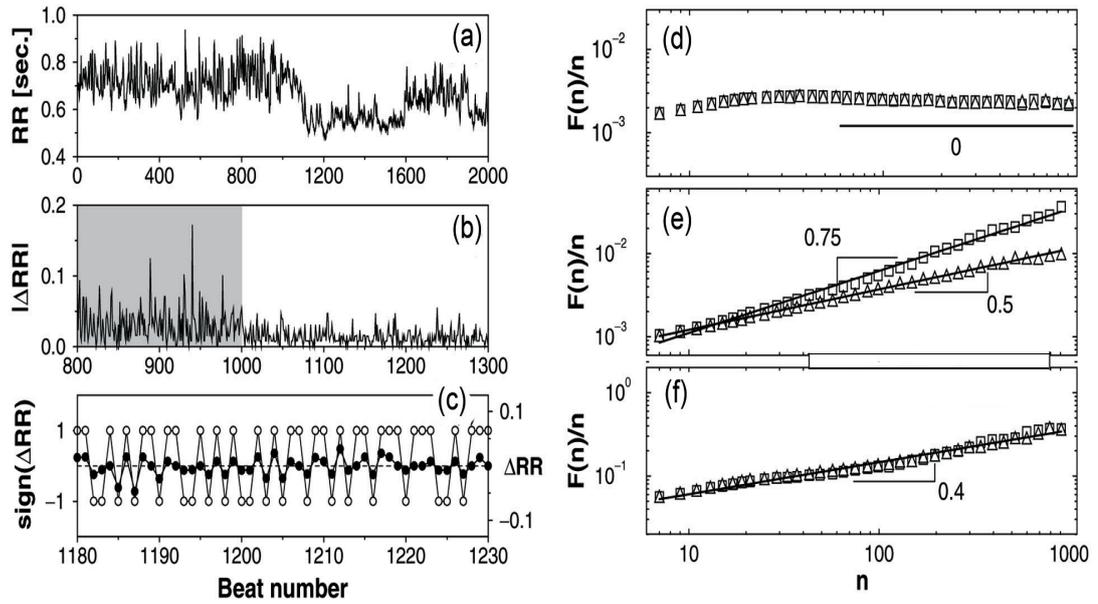


Figura 3.4: Correlações de longo alcance na dinâmica dos batimentos cardíacos. Fonte: ref. [92]. **(a)** Intervalos de tempo entre batimentos cardíacos sucessivos de um indivíduo saudável, $RR(i)$, sendo i o número do batimento (*Beat number*). **(b)** Magnitude dos incrementos, $|\Delta RR(i)|$, com $\Delta RR(i) = RR(i + 1) - RR(i)$. **(c)** Sinal dos incrementos, $\text{sign}[\Delta RR(i)]$, fornecendo +1 se $\Delta RR(i) > 0$ e -1 se $\Delta RR(i) < 0$. Os próprios incrementos $\Delta RR(i)$ são também representados (símbolos preenchidos). **(d)** Função de flutuação $F(n)$ (obtida pelo método *detrended fluctuation analysis* (DFA)) dividida pela escala de tempo n para a série dos incrementos $\Delta RR(i)$. **(e)** o mesmo que em (d), mas para a série das magnitudes $|\Delta RR(i)|$; **(f)** o mesmo que em (d), mas para a série dos sinais $\text{sign}[\Delta RR(i)]$. Observe que $F(n)/n \sim n^h$, com $h \simeq 0$ para a série dos incrementos, $h \simeq 0,75$ para a série das magnitudes e $h \simeq 0,4$ para a série dos sinais.

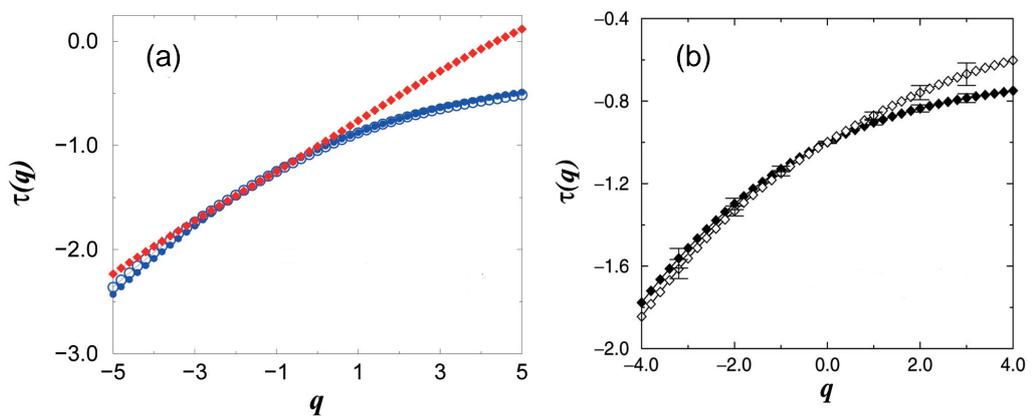


Figura 3.5: Multifractalidade na dinâmica dos batimentos cardíacos. Fonte: refs. [93, 94]. **(a)** Espectro multifractal $\tau(q)$ para *i*) um indivíduo saudável acordado (círculos azuis abertos); *ii*) um indivíduo saudável dormindo (círculos azuis fechados); e *iii*) um indivíduo doente (círculos vermelhos fechados). A curvatura de $\tau(q)$ indica multifractalidade na série “saudável” enquanto que o comportamento praticamente linear de $\tau(q)$ indica monofractalidade na série “doente”. **(b)** Espectro multifractal $\tau(q)$ para um indivíduo saudável na sua rotina normal (losangos) e um indivíduo saudável em repouso controlado (losangos preenchidos). Em ambos os casos, as séries apresentam multifractalidade.

crecem com a escala de tempo como leis de potência, com expoentes $\delta \simeq 0$ para o indivíduo saudável e $\delta \simeq 0,5$ para o indivíduo doente (fig. 3.3c). Esses resultados indicam a presença de correlações de longo alcance na série “saudável” e ausência de correlações de longo alcance na série “doente” [91].

Ademais, a distribuição dos incrementos consecutivos (escala de tempo $n = 1$) deste sinal fisiológico apresenta comportamento não-Gaussiano. Conforme ilustrado na fig. 3.3d, as distribuições para as séries “saudável” e doente” são bem descritas por uma distribuição simétrica de Lévy (veja apêndice A), com índice $\psi = 1,7$ [91]. As flutuações na série dos batimentos cardíacos continuam sendo não-Gaussianas para grandes escalas de tempo. Elas mantêm a sua forma funcional para diferentes escalas de tempo, exibindo invariância de escala [95].

A série dos incrementos consecutivos nos intervalos de tempo entre batimentos cardíacos pode ser decomposta em duas sub-séries: a série das magnitudes e a série dos sinais [92]. As figs. 3.4a, 3.4b e 3.4c ilustram este procedimento. Como pode ser visto nas figs. 3.4d, 3.4e e 3.4f, as flutuações crescem aproximadamente como leis de potência, com expoentes $h \simeq 0$ para a série dos incrementos, $h \simeq 0,75$ para a série das magnitudes e $h \simeq 0,4$ para a série dos sinais. Observe que o expoente $h \simeq 0$ na fig. 3.4d é consistente com o resultado $\delta \simeq 0$ visto na fig. 3.3c para o indivíduo saudável. Tanto a série das magnitudes como a série dos sinais fornecem resultados (expoentes) distintos para indivíduos saudáveis e doentes [92]. Os expoentes h também assumem valores distintos em diferentes estágios do sono e diferentes níveis de atividade física [97-99].

A série dos intervalos de tempo entre batimentos cardíacos também exhibe estrutura multifractal [93, 94]. Na fig. 3.5a, vemos o espectro multifractal $\tau(q)$ para um indivíduo saudável acordado, um indivíduo saudável durante o sono e um indivíduo doente. Observe que $\tau(q)$ é uma função não linear de q para as séries “saudáveis”, enquanto que $\tau(q)$ é praticamente linear para a série “doente” [93]. Ademais, na fig. 3.5c vemos o espectro multifractal $\tau(q)$ para um indivíduo saudável em duas situações distintas: atividade normal e repouso controlado. Em ambos os casos, a série exhibe $\tau(q)$ não linear [94]. Para uma revisão mais completa sobre o tema veja, por exemplo, a ref. [100].

A seguir, consideraremos dois tipos de sinais complexos cuja dinâmica apresenta intrigantes similaridades com a dinâmica dos batimentos cardíacos.

3.2.2 Sistema hidrodinâmico

Torneiras gotejantes são sistemas hidrodinâmicos aparentemente simples, mas que exibem comportamento complexo. Considere a série dos intervalos de tempo entre gotas sucessivas numa torneira gotejante. Esta série exhibe anti-correlações temporais de longo alcance e distribuição não-Gaussiana, conforme apresentado na ref. [101] e ilustrado na fig. 3.6.

Especificamente, as flutuações crescem com a escala de tempo como leis de potência, com expoentes $\delta \simeq 0$ (fig. 3.6a), indicando que esta série apresenta anti-correlações de longo alcance. Ademais, a distribuição dos incrementos consecutivos apresenta comportamento não-Gaussiano, sendo bem descrita por uma distribuição de Lévy com índice $\psi = 1,7$ e $\rho > 0$ (fig. 3.6b).

Há uma intrigante analogia entre a dinâmica dos intervalos de tempo entre gotas numa torneira gotejante e a dinâmica dos batimentos cardíacos[101]. Uma comparação dos resultados apresentados nas figs. 3.6 e 3.3 mostra que ambos os sistemas exibem leis empíricas quantitativamente similares. Especificamente, em ambos os casos, as séries dos incrementos apresentam anti-correlações de longo alcance, caracterizadas pelo mesmo expoente $\delta \simeq 0$. Além disso, os incrementos consecutivos em ambos os sistemas seguem uma distribuição de Lévy, com o mesmo índice $\psi \simeq 1,7$.

Uma possível explicação para tais similaridades entre sistemas tão diferentes é a presença de competição em ambos os sistemas[101]. Flutuações nos intervalos de tempo entre batimentos cardíacos resultam da competição entre estímulos do sistema nervoso autônomo[91, 92]. Analogamente, numa torneira gotejante há competição entre forças que criam tensão superficial e a força da gravidade[101].

É bem conhecido que diversos sistemas exibem anti-correlações de longo alcance com expoente $\delta \simeq 0$. Entretanto, o fato de que a distribuição das flutuações segue a mesma distribuição e com o mesmo índice reforça a analogia comentada. Estas similaridades quantitativas são um exemplo de comportamento universal em sistemas complexos. Com base nestas similaridades, é possível considerar a possibilidade de que algum tipo de mecanismo em comum atue na dinâmica de ambos os sistemas[101].

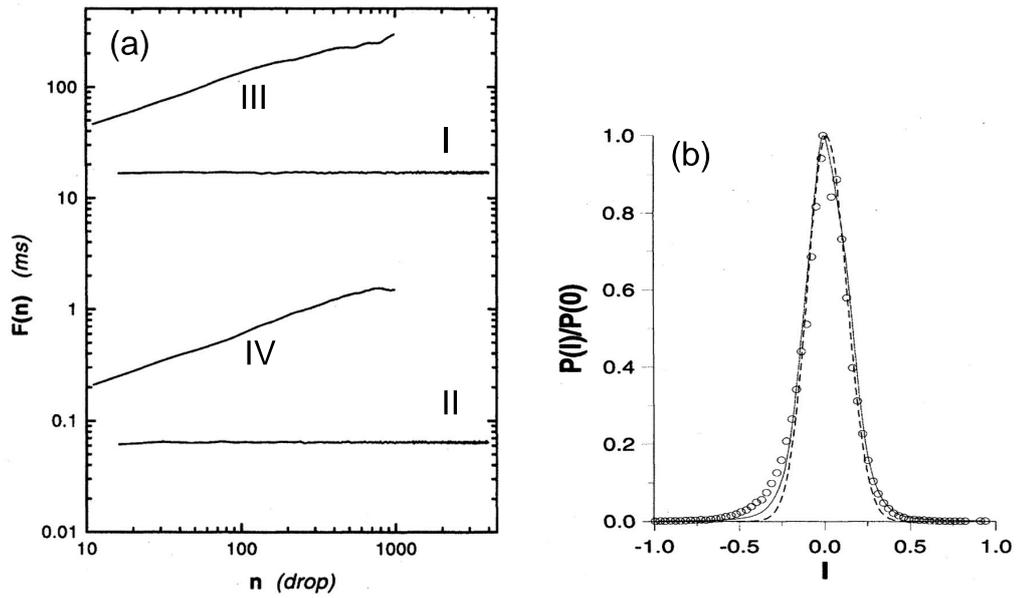


Figura 3.6: Correlações de longo alcance e distribuição não-Gaussiana na dinâmica da torneira gotejante. Fonte: ref. [101]. **(a)** Função de flutuação $F(n)$, definida na ref. [101], para as seguintes séries: I) série dos intervalos de tempo entre gotas sucessivas, para uma taxa de fluxo de 25 gotas por segundo; II) série dos intervalos de tempo entre gotas sucessivas, para uma taxa de fluxo de 40 gotas por segundo; III) versão embaralhada da série em I); IV) versão embaralhada da série em II). A função $F(n)$ para as versões embaralhadas referem-se a médias sobre diversas realizações. Observe que $F(n) \sim n^\delta$, com $\delta \simeq 0$ para as séries originais e $\delta \simeq 0,5$ para as versões embaralhadas. **(d)** Distribuição dos incrementos sucessivos, $p(I)/p(0)$. A linha sólida representa uma distribuição de Lévy com $\psi = 1,7$ e a linha descontínua corresponde a uma distribuição Gaussiana ($\psi = 2,0$).

3.2.3 Tráfego na Internet

A dinâmica de transferência de informações na Internet também exhibe similaridades com a dinâmica cardíaca. Considere resultados descritos na ref. [102] e ilustrados na fig. 3.7.

As séries analisadas fornecem o número total de funcionários de um certo grupo de laboratórios (laboratórios NTT, de Tóquio) conectados com a Internet a cada instante de tempo (a cada segundo). As propriedades de correlação de tais séries são significativamente diferentes para períodos da noite, quando o tráfego na internet está descongestionado e períodos do dia em que o tráfego está congestionado.

Como pode ser visto na fig. 3.7a, as flutuações na série em questão crescem como leis de potência, com expoentes $h \simeq 1$ para séries no período da noite (tráfego descongestionado) e $h \simeq 1,3$ para séries no período do dia (tráfego congestionado). Esses resultados indicam a presença de correlações de longo alcance na série “descongestionada” e uma significativa perda de correlações de longo alcance para a série “congestionada”.

Algo análogo ocorre com os intervalos de tempo entre batimentos cardíacos, conforme ilustrado na fig. 3.7b. As flutuações na série dos intervalos de tempo entre batimentos cardíacos são caracterizadas por expoentes $h \simeq 1$ (o que corresponde a $h \simeq 0$ para a série dos incrementos) para indivíduos saudáveis e $h \simeq 1,3$ para indivíduos doentes[102] (comparare as figs. 3.7a e 3.7b). Este é mais um exemplo de comportamento universal na dinâmica de sistemas complexos. Estas similaridades sugerem que a rede de comunicação da Internet pode ser controlada por mecanismos similares aos que controlam a rede de comunicação do sistema nervoso autônomo[102].

Seguindo as linhas de pesquisa descritas neste capítulo, passaremos a analisar os seguintes sistemas: circulação de revistas, citações de periódicos científicos, atividades religiosas e um sinal geomagnético. Iniciaremos com a circulação de revistas, no próximo capítulo.

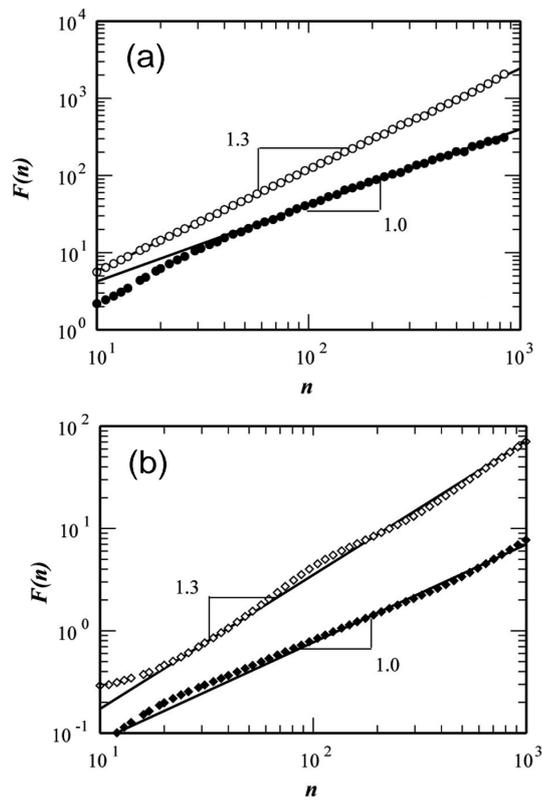


Figura 3.7: Correlações de longo alcance no tráfego da internet. Fonte: ref. [102]. **(a)** Função de flutuação $F(n)$ (obtida pelo método *detrended fluctuation analysis* (DFA)) para a série das conexões para período do dia (círculos abertos) e período da noite (círculos fechados). Observe que $F(n) \sim n^h$, com $h \simeq 1$ para a série “descongestionada” e $h \simeq 1,3$ para a série “congestionada”. **(b)** Função de flutuação $F(n)$ (obtida pelo método *detrended fluctuation analysis* (DFA)) para a série dos batimentos cardíacos para indivíduo doente (círculos abertos) e indivíduo saudável (círculos fechados). Observe que $F(n) \sim n^h$, com $h \simeq 1$ para a série “saudável” e $h \simeq 1,3$ para a série “doente”. Note a similaridade quantitativa entre os expoentes para dinâmica cardíaca e tráfego na Internet.

Capítulo 4

Circulação de revistas e jornais

Investigaremos aqui propriedades estatísticas da circulação de revistas e jornais[20]. Especificamente, obteremos *i*) a distribuição da circulação; *ii*) a distribuição de suas taxas semestrais de crescimento; e *iii*) o desvio padrão das taxas de crescimento em termos da circulação. Além disso, compararemos a dinâmica de crescimento da circulação de revistas com a dinâmica de crescimento do tamanho de firmas. Proporemos também um modelo estocástico, inspirado num modelo para o crescimento de firmas, e compararemos suas predições com os resultados empíricos para a circulação de revistas.

4.1 Introdução

Conforme exposto no capítulo 3, a dinâmica de crescimento de atividades econômicas exibe padrões universais. Em particular, no que se refere ao tamanho de firmas, várias medidas de tamanho têm sido consideradas - desde o total das vendas ao valor total das propriedades e equipamentos da firma em questão. Neste capítulo, consideraremos um tipo particular de organização econômica e investigaremos uma medida adicional de tamanho: a circulação de revistas e jornais.

A circulação de publicações impressas tem sido objeto de estudo principalmente em marketing e economia [103-106]. O termo circulação refere-se basicamente ao número médio de cópias de uma publicação que são distribuídas ao público. A circulação é definida como a razão entre o número total de cópias distribuídas durante um certo período pelo número de edições da publicação no período considerado. O número total de cópias inclui os números avulsos vendidos e os exemplares de assinantes (circulação paga) bem

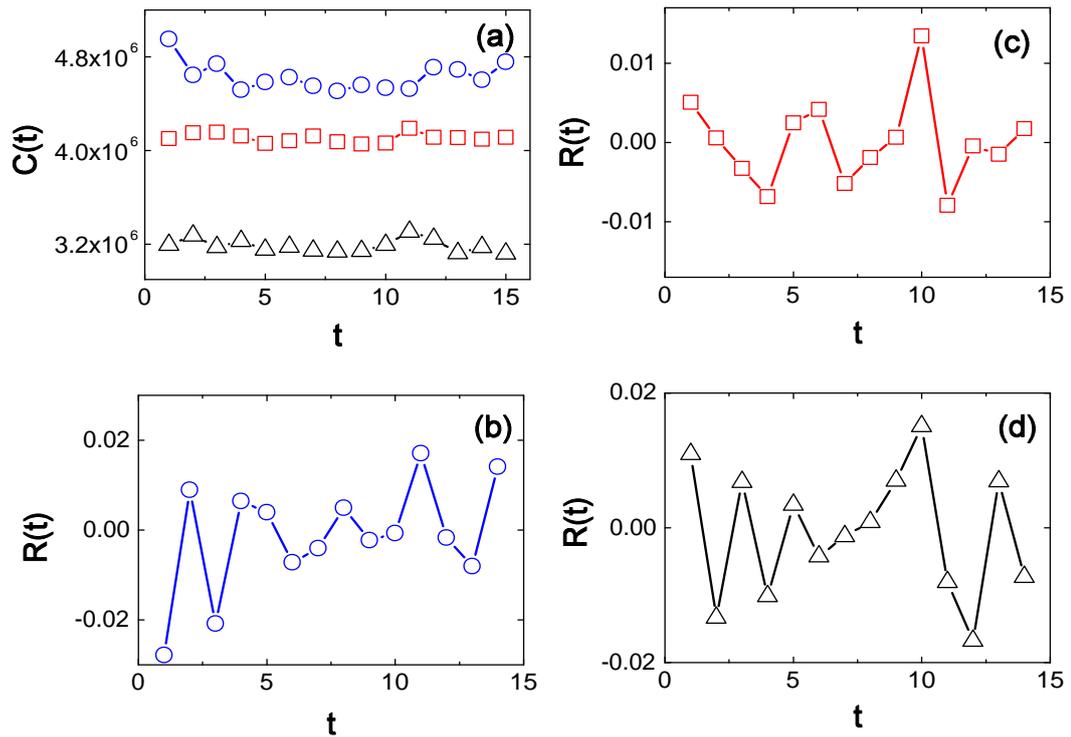


Figura 4.1: **(a)** Evolução temporal da circulação, $C(t)$, das revistas *Good House Keeping* (círculos), *Time* (quadrados) e *Newsweek* (triângulos) num período de 15 semestres entre 1996-2003. **(b)** Evolução temporal das taxas de crescimento $R(t) = \log[C(t+1)/C(t)]$ calculadas para a revista *Good House Keeping*, no período 1996-2003. **(c)** O mesmo que em (b), mas para a revista *Time*. **(d)** O mesmo que em (b), mas para a revista *Newsweek*.

como cópias direcionadas sem custo a locais ou indivíduos específicos (circulação não paga). A circulação costuma ser divulgada, inicialmente, como uma estimativa do editor quanto ao número de cópias a serem distribuídas num período específico. Empresas que prestam serviços de auditoria, tais como *Audit Bureau of Circulations* (ABC)[107] and *BPA Worldwide* (BPA)[108], verificam a circulação real da publicação e divulgam seus resultados.

Utilizamos dados que incluem a circulação das 100 maiores revistas do mundo, das 570 maiores revistas nos Estados Unidos, das 727 maiores revistas no Reino Unido e dos 100 maiores jornais nos Estados Unidos. Incluem também a circulação de 82 revistas nos Estados Unidos, num período de 15 semestres de 1996 a 2003. Para ilustrar, na fig. 4.1 vemos a evolução temporal da circulação e das taxas de crescimento de três revistas no período 1996-2003. No que segue, descreveremos os resultados obtidos da análise desses dados.

4.2 Distribuição da circulação

Iniciamos calculando a distribuição acumulada da circulação, $p_c(C)$, para as 100 maiores revistas nos Estados Unidos, no Reino Unido e em âmbito mundial e para os 100 maiores jornais nos Estados Unidos. Conforme vemos na fig. 4.2a, a distribuição acumulada apresenta um comportamento assintótico do tipo lei de potência,

$$p_c(C) \sim C^{-\eta}, \quad (4.1)$$

com $\eta \simeq 1,5$, para todos os conjuntos de dados considerados. Este resultado indica $p(C) \sim C^{-\alpha}$, com $\alpha \simeq 2,5$, sendo $p(C)$ a distribuição da circulação.

A distribuição acumulada $p_c(C)$ desvia-se do comportamento do tipo lei de potência para valores menores de C , como pode ser visto na fig. 4.2b. Nesta figura vemos $p_c(C)$ calculada para 570 revistas nos Estados Unidos e 727 revistas no Reino Unido, em 2004, em comparação com uma distribuição q -exponencial

$$p_c(C) = a' \left[1 - (1 - q') \frac{C}{b'} \right]^{\frac{-1}{q'-1}}, \quad (4.2)$$

com $q' = 1,66$. Observe que $p_c(C) \sim C^{1/(1-q')} = C^{-1,51}$ para grandes valores de C , reproduzindo o comportamento assintótico visto na fig. 4.2a. Este resultado indica que

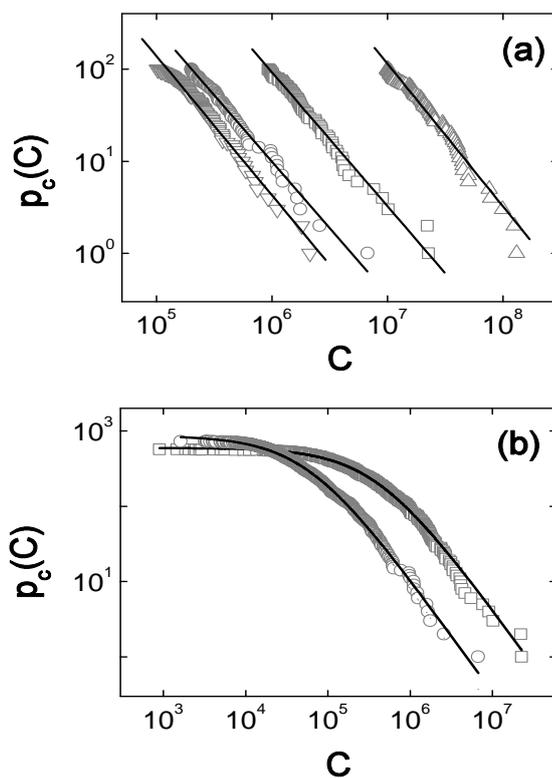


Figura 4.2: Distribuição da circulação de revistas e jornais. **(a)** Distribuição acumulada, $p_c(C)$, para os 100 maiores jornais nos EUA em 2003 (triângulos para baixo), as 100 maiores revistas no Reino Unido em 2004 (círculos), as 100 maiores revistas nos EUA em 2004 (quadrados) e as 100 maiores revistas do mundo em 2002 (triângulos para cima: dados multiplicados por 10). As linhas retas representam leis de potência com expoentes $\eta \simeq 1,50 \pm 0,05$. Isso indica que $p(C) \sim C^{-\alpha}$, com $\alpha \simeq 2,5$. **(b)** Distribuição acumulada, $p_c(C)$, calculada para 570 revistas nos EUA em 2004 (quadrados) e de 727 revistas no Reino Unido em 2005 (círculos). As linhas sólidas representam a distribuição q -exponencial, dada pela eq. (4.2), com $q' = 1,66$, $b' = 2,52 \cdot 10^5$ e $a' = 5,95 \cdot 10^2$ para os EUA e $q' = 1,66$, $b' = 3,72 \cdot 10^4$, e $a' = 8,62 \cdot 10^2$ para o Reino Unido.

a distribuição correspondente $p(C)$ será descrita por uma q -exponencial com $q \simeq 1,4$ (veja apêndice A).

Curiosamente, o comportamento observado na fig. 4.2b, para a circulação de revistas, é quantitativamente similar ao que se observa na distribuição de citações de artigos científicos[35], população de cidades[37] e número de votos entre candidatos numa eleição[38]. Especificamente, em todos estes casos, a distribuição da variável em questão tem sido descrita como sendo q -exponencial, com $q \simeq 1,4$ [35, 37, 38].

4.3 Distribuição das taxas de crescimento

Investigamos também propriedades estatísticas das taxas de crescimento da circulação de revistas. Seguindo a definição usada no estudo de firmas, definimos taxa de crescimento como

$$R(t) = \log \left[\frac{C(t+1)}{C(t)} \right], \quad (4.3)$$

com $C(t)$ e $C(t+1)$ representando a circulação de uma dada revista nos semestres t e $t+1$, respectivamente.

Calculamos as taxas de crescimento $R(t)$ para 82 revistas americanas, com grande circulação, num período de 15 semestres entre 1996 e 2003. Com base nesses dados, obtemos a densidade de probabilidade condicional, $p(R|C)$, das taxas de crescimento da circulação de revistas. Conforme ilustrado na fig. 4.3a, $p(R|C)$ exhibe um comportamento não-Gaussiano, sendo bem ajustada pela função

$$p(R|C) = \frac{1}{\sqrt{2}\sigma(C)} \exp \left(-\frac{\sqrt{2}|R - \mu|}{\sigma(C)} \right). \quad (4.4)$$

Nesta equação, μ é a média e σ é o desvio padrão calculados sobre todas as revistas do grupo considerado. Para o mesmo grupo de revistas, calculamos as taxas de crescimento normalizadas, definidas como

$$r(t) = \frac{R(t) - \mu}{\sigma}. \quad (4.5)$$

Na fig. 4.3b, vemos a distribuição das taxas de crescimento normalizadas $p(r)$ em comparação com a função

$$p(r) = \frac{1}{\sqrt{2}} \exp(-\sqrt{2}|r|). \quad (4.6)$$

Calculamos também o desvio padrão σ das taxas de crescimento R em termos da

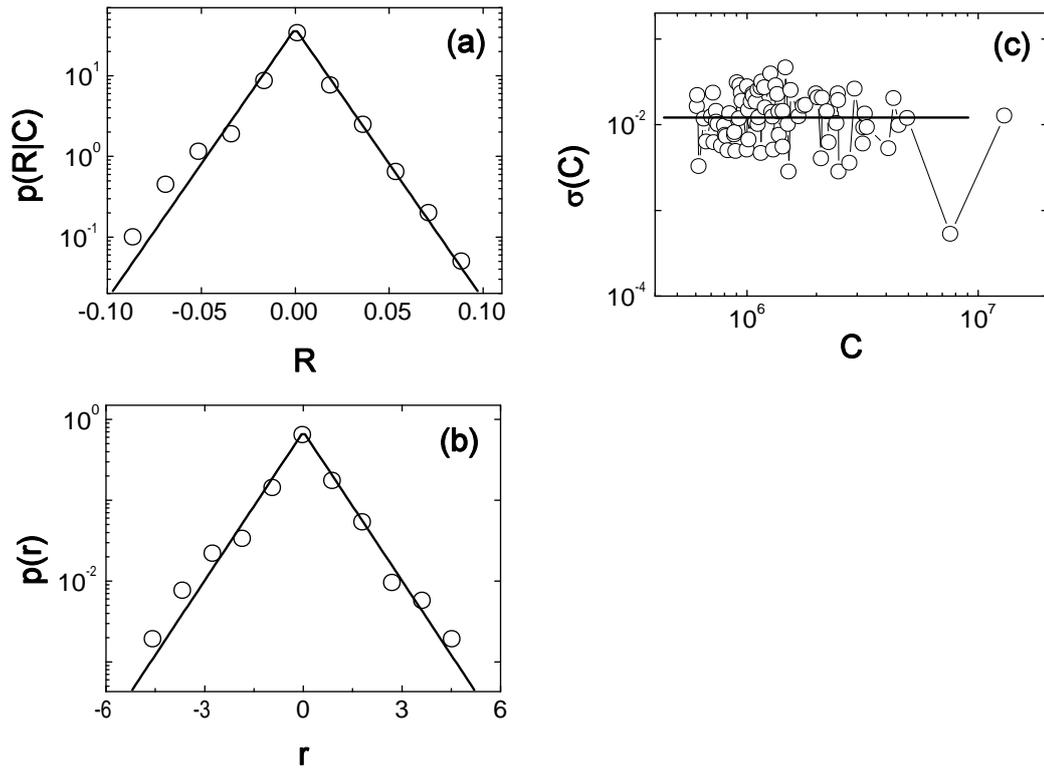


Figura 4.3: Dinâmica de crescimento da circulação de revistas: as mais lidas. **(a)** Densidade de probabilidade condicional $p(R|C)$ para um grupo de 82 revistas nos Estados Unidos, com $6,08 \cdot 10^5 < C < 1,30 \cdot 10^7$, num período de 15 semestres (de 1996 a 2003). A linha sólida representa a eq. (4.4), com μ e σ calculados diretamente a partir dos dados. **(b)** Densidade de probabilidade $p(r)$ para as mesmas revistas consideradas em (a). A linha sólida representa a eq. (4.6). **(c)** Desvio padrão σ (calculado para cada uma das 82 revistas consideradas) das taxas de crescimento R em função da circulação inicial C . A linha horizontal representa $\sigma(C) \sim C^{-\beta}$, com $\beta = 0$.

circulação C para cada uma das revistas consideradas. A fig. 4.3c indica que

$$\sigma(C) \propto C^{-\beta}, \quad (4.7)$$

com $\beta \simeq 0$, indicando que σ (calculado para cada revista) é praticamente independente de C para grandes valores de C . A seguir, consideraremos características universais na dinâmica da circulação.

4.4 Características universais

Os principais resultados considerados nas duas seções anteriores, para revistas com grande circulação, foram: *i*) a distribuição acumulada $p_c(C)$ exibe comportamento assintótico do tipo lei de potência, com expoente $\eta \simeq 1,5$, indicando que $p(C) \sim C^{-\alpha}$, com $\alpha \simeq 2,5$; *ii*) a distribuição das taxas de crescimento $p(R|C)$ exibe decaimento exponencial; e *iii*) o desvio padrão $\sigma(C)$ é praticamente independente de C , exibindo expoente $\beta \simeq 0$. Esses resultados são similares aos observados na dinâmica de grandes firmas (veja tab. 4.1).

Tabela 4.1 - Dinâmica de crescimento das revistas mais lidas e das firmas que mais vendem (capítulo 3).

	X	$p(X)$	α	$\sigma(X)$	β	$p(R X)$
Revistas	circulação	$\sim X^{-\alpha}$	2,5	$\sim X^{-\beta}$	0	exponencial
Firmas	vendas	$\sim X^{-\alpha}$	2,0	$\sim X^{-\beta}$	0	exponencial

Motivados por estas similaridades, sugerimos uma possível interpretação dos resultados resumidos na tab. 4.1 em termos de um processo estocástico inspirado no modelo de Gibrat para o crescimento de firmas. Esse é o assunto da próxima seção.

4.5 Modelo estocástico I

Vamos supor que a circulação $C(t)$ de uma dada revista obedeça ao processo estocástico multiplicativo

$$C(t+1) - C(t) = \lambda(t)C(t), \quad (4.8)$$

com $C \geq C_m$ ($C = \max\{C, C_m\}$), sendo C_m uma constante positiva e diferente de zero. Suponha também que $\lambda(t)$ seja um processo multiplicativo do tipo

$$\lambda(t) = [a_0 + a_1\lambda(t-1)]\epsilon_\nu(t), \quad (4.9)$$

sendo $a_0 \ll 1$ e a_1 constantes positivas. O termo $\epsilon_\nu(t)$ refere-se a um número aleatório dado pela relação

$$\epsilon_\nu(t) = \begin{cases} -|N(0, 1)| & p = 0.5 \\ |N(0, \nu)| & p = 0.5, \end{cases} \quad (4.10)$$

em que o termo $N(0, \nu)$ refere-se a um número aleatório cuja distribuição é Gaussiana, com média zero e variância ν . Na eq. (4.10), $p = 0, 5$ significa “com probabilidade 0,5”. Se $\nu = 1$, a eq. (4.10) descreverá simplesmente um ruído Gaussiano, $\epsilon_1(t) = N(0, 1)$. Por outro lado, se $\nu \neq 1$, o termo $\epsilon_\nu(t)$ passará a exibir uma distribuição gaussiana assimétrica. Note que o processo estocástico definido pelas eqs. (4.8), (4.9) e (4.10) recupera o modelo de Gibrat (na ausência do corte C_m) quando $\nu = 1$ e $a_1 = 0$.

A introdução da restrição $C \geq C_m$ é razoável visto que ela mantém positivos os valores da circulação C . A presença desta barreira refletora é fundamental para que o processo gere distribuições $p_c(C) \propto C^{-\eta}$. Verificamos que valores particulares de C_m e de $C_0 \geq C_m$ não influenciam o valor do expoente η da distribuição. Verificamos também que este processo estocástico é consistente com a equação $\sigma(C) \propto C^{-\beta}$, com $\beta \simeq 0$, indicando que o desvio padrão das taxas de crescimento é independente do tamanho. De modo similar, verificamos que o valor médio das taxas de crescimento também exibe um comportamento geral independente do tamanho, principalmente para valores de C distantes da barreira refletora definida por C_m .

Para investigar o efeito dos principais parâmetros, a_1 e ν , sobre certas propriedades estatísticas de C e $R = \log[C(t+1)/C(t)]$, realizamos uma série de simulações numéricas com $0 \leq a_1 \leq 1$ e $0.8 \leq \nu \leq 1$. Os principais resultados dessas simulações podem ser vistos na fig. 4.4.

O parâmetro a_1 controla o aspecto autoregressivo do processo, produzindo um efeito

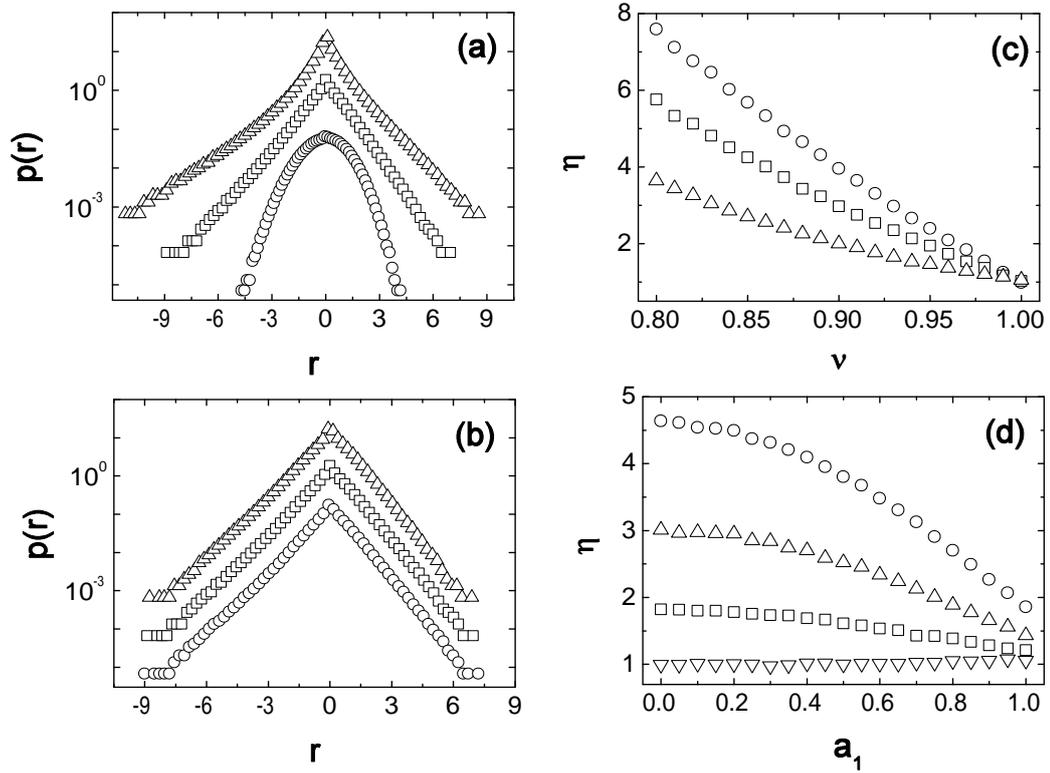


Figura 4.4: Simulação (média sobre diversas realizações) do processo estocástico definido pelas eqs. (4.8), (4.9) e (4.10). Consideramos 1000 revistas, todas com circulação inicial $C_0 = 10^3$ e com $C_m = 1$. Fixamos $a_0 = 0.03$ e empregamos diversos valores de a_1 e ν , considerando 150 passos depois de um regime inicial. **(a)** Densidade de probabilidade $p(r)$ para $\nu = 0,97$ e $a_1 = 0$ (círculos), $a_1 = 0,6$ (quadrados) e $a_1 = 0,9$ (triângulos). **(b)** Densidade de probabilidade $p(r)$, para $a_1 = 0.6$ e $\nu = 1$ (círculos), $\nu = 0,97$ (quadrados) e $\nu = 0.93$ (triângulos). Em (a) e (b) as curvas foram deslocadas na vertical para melhor visualização. **(c)** O expoente η , da relação $p_c(C) \sim C^{-\eta}$, em termos do parâmetro ν , para $a_1 = 0$ (círculos), $a_1 = 0,6$ (quadrados) e $a_1 = 0,9$ (triângulos). **(d)** O expoente η em termos do parâmetro a_1 , para $\nu = 1$ (triângulos para baixo), $\nu = 0,97$ (quadrados), $\nu = 0,93$ (triângulos para cima) e $\nu = 0,88$ (círculos).

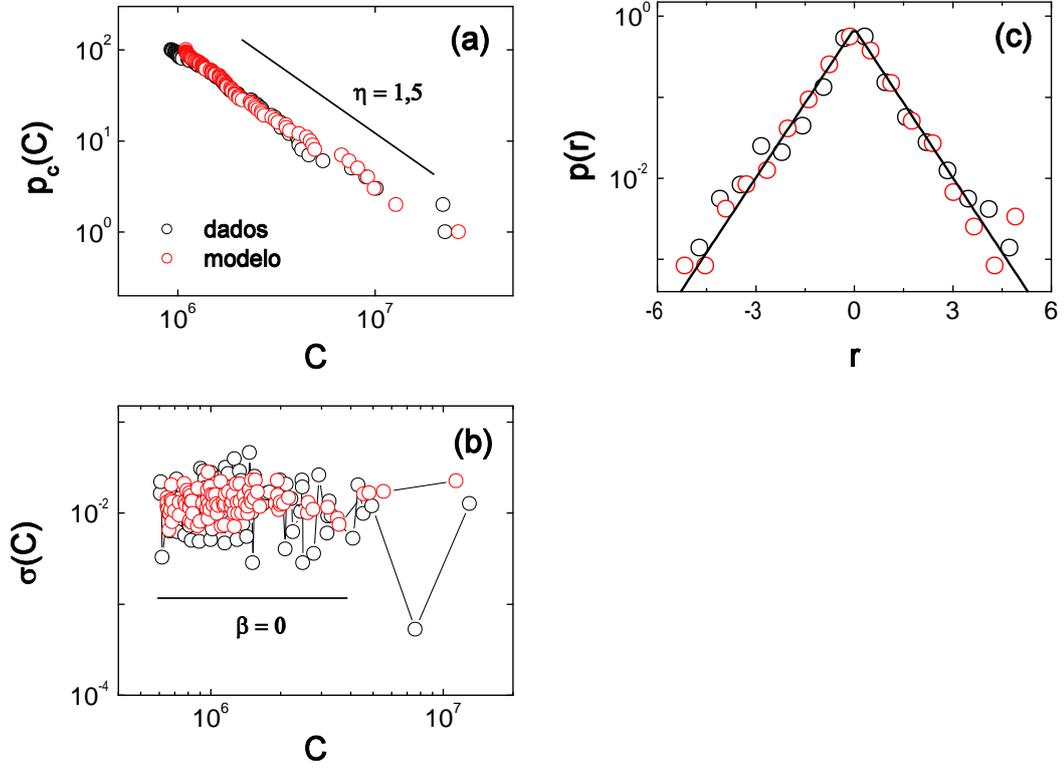


Figura 4.5: Realização típica do processo definido pelas eqs. (4.8), (4.9) e (4.10) em comparação com resultados empíricos para a circulação de revistas. Na simulação, usamos 1000 revistas com $C_0 = 1$, $C_m = 1$ e $t = 20$ passos depois de descartar os passos iniciais. Fixamos $a_0 = 0.03$, $a_1 = 0,6$ e $\nu = 0,97$. (a) Distribuição acumulada, $p_c(C)$, para as 100 maiores revistas (círculos vermelhos). Os dados empíricos (círculos pretos) referem-se às 100 maiores revistas nos EUA em 2004 (fig. 4.2a). A linha sólida representa uma lei de potência com expoente $\eta = 1,5$. (b) Desvio padrão σ (calculado para cada revista) das taxas de crescimento R em função de C *i*) para as 100 maiores revistas na simulação (círculos preenchidos); e *ii*) para os dados empíricos mostrados na fig. 4.3b (círculos). A linha horizontal representa uma lei de potência com expoente $\beta = 0$. Os resultados da simulação em (a) e (b) foram convenientemente deslocados horizontalmente para melhor comparação. (c) Densidade de probabilidade $p(r)$ das taxas de crescimento normalizadas r *i*) para as 100 maiores revistas na simulação (círculos fechados); e *ii*) para os dados empíricos mostrados na fig. 4.3a (círculos). A linha reta representa a eq. (4.6).

de memória no sistema. Este parâmetro regula a intensidade com que a taxa (atual) de crescimento de uma revista depende da taxa de crescimento anterior. Como vemos na fig. 4.4a, o parâmetro a_1 está intimamente relacionado à forma da distribuição das taxas de crescimento. A distribuição é Gaussiana se $a_1 = 0$, desviando-se da Gaussiana à medida que a_1 cresce. Em contraste, como observado na fig. 4.4b, a forma da distribuição das taxas de crescimento não apresenta dependência significativa com o parâmetro ν , principalmente para ν perto da unidade.

Por sua vez, o parâmetro ν controla o aspecto assimétrico do ruído $\epsilon_\nu(t)$ (veja eq. (4.10)). Valores de ν menores que a unidade restringem os grandes valores positivos das taxas de crescimento, favorecendo os grandes decréscimos. Ambos os parâmetros, ν e a_1 , são responsáveis por mudanças no expoente η da distribuição $p_c(C)$. Nas figs. 4.4c e 4.4d vemos o expoente η em termos dos parâmetros ν e a_1 , respectivamente. Observe que o expoente η é extremamente sensível ao parâmetro ν . Note também que há uma competição entre ν e a_1 para dar um valor específico de η . À medida que ν se afasta da unidade, a assimetria aumenta e η cresce. Por outro lado, o valor de η diminui à medida que a_1 aumenta.

Na fig. 4.5, comparamos as previsões deste processo estocástico para um dado conjunto de parâmetros com os principais resultados empíricos apresentados neste capítulo. Fixando $a_1 = 0,6$, escolhemos $\nu = 0,97$ de tal forma que gere um expoente $\eta \simeq 1,5$ (figs. 4.4c e 4.4d). Observe a grande sensibilidade do expoente η ao parâmetro ν : no caso mencionado acima, uma mudança de apenas 0,03 em ν causa uma mudança de 0,5 em η . Uma realização típica deste processo para o conjunto escolhido de parâmetros pode ser vista na fig. 4.5 em comparação com resultados empíricos. Observe que o modelo reproduz satisfatoriamente os principais aspectos do comportamento observado.

Capítulo 5

Citações de periódicos científicos

Neste capítulo, investigaremos propriedades estatísticas de uma medida das citações de periódicos científicos - o fator de impacto[21]. Especificamente, calcularemos *i*) a distribuição do fator de impacto; *ii*) a distribuição de suas taxas anuais de crescimento; e *iii*) o desvio padrão das taxas de crescimento em termos do fator de impacto. Além disso, discutiremos aspectos universais observados na dinâmica desta medida de atividade de pesquisa. Proporemos também um modelo para a dinâmica de crescimento de organizações e compararemos suas previsões com os principais resultados descritos no capítulo.

5.1 Introdução

No capítulo 3, vimos que a dinâmica de crescimento do tamanho de universidades é governada por leis universais. As medidas de tamanho de universidades incluem fundos de pesquisa, número de patentes e número de artigos publicados anualmente. Neste capítulo, analisamos outro tipo de atividade de pesquisa - as citações de periódicos científicos. Especificamente, investigamos a dinâmica de crescimento do fator de impacto dos periódicos.

O fator de impacto de periódicos científicos foi proposto na década de 1950[109, 110]. É definido como a razão entre o número de citações que um periódico recebe no ano corrente, devido a artigos publicados nos dois anos anteriores, pelo número de artigos publicados nos dois anos anteriores. Por exemplo, para calcular o fator de impacto de um periódico relativo ao ano 2006, dividimos as citações que ele recebeu em 2006, devido

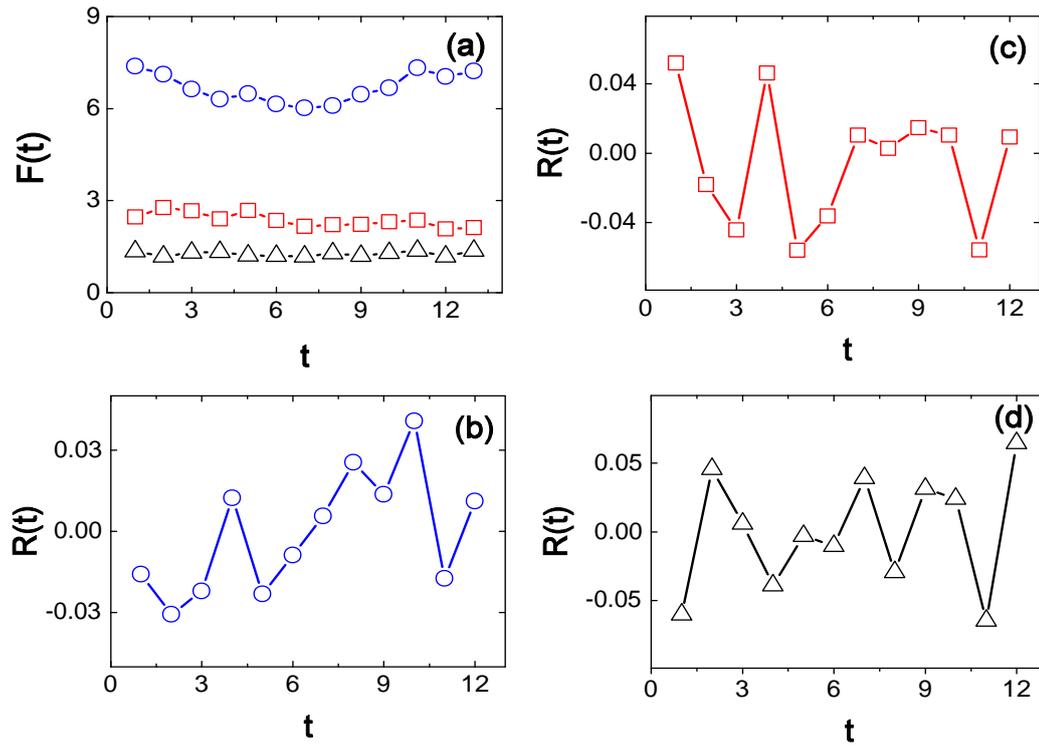


Figura 5.1: **(a)** Evolução temporal do fator de impacto, $F(t)$, dos periódicos *Physical Review Letters* (círculos), *Europhysics Letters* (quadrados) e *Physica A* (triângulos) num período de 13 anos (1992-2004). **(b)** Evolução temporal das taxas de crescimento $R(t) = \log[F(t + 1)/F(t)]$ calculadas para o periódico *Physical Review Letters*, no período 1992-2004. **(c)** O mesmo que em (b), mas para o periódico *Europhysics Letters*. **(d)** O mesmo que em (b), mas para o periódico *Physica A*.

a artigos publicados em 2004 e 2005, pelo número de artigos publicados em 2004 e 2005. O fator de impacto é calculado anualmente pelo *Institute for Scientific Information* (ISI) e publicado no *Journal Citation Report* (JCR)[111]. Em anos recentes, esta medida tem sido amplamente utilizada, por exemplo, na administração de coleções de periódicos em bibliotecas e até na avaliação acadêmica.

Utilizamos um conjunto de dados que inclui o fator de impacto de mais de 5000 periódicos em vários anos. Incluem também o fator de impacto de 2910 periódicos num período de 13 anos: de 1992 a 2004. Vemos na fig. 5.1 a evolução temporal do fator de impacto e das taxas de crescimento de três periódicos neste período. Segue uma descrição dos resultados obtidos da análise destes dados.

5.2 Distribuição do fator de impacto

Começamos calculando a distribuição do fator de impacto de 5912 periódicos em 2004. Na fig. 5.2a, vemos $p(\log F)$ em comparação com uma distribuição Gaussiana. Observe que a distribuição empírica $p(\log F)$ se desvia do comportamento Gaussiano para pequenos e grandes valores do fator de impacto, indicando que $p(F)$ não pode ser satisfatoriamente descrita por uma distribuição log-normal. De fato, um estudo prévio já havia indicado que $p(F)$ comporta-se como lei de potência para grandes valores de F [112].

A distribuição $p(F)$ exibe comportamento do tipo lei de potência para grandes e pequenos valores de F . Conforme ilustrado na fig. 5.2b,

$$p(F) \sim F^{-\alpha}, \quad (5.1)$$

com $\alpha \simeq 2,7$ para F grande e $\alpha \simeq -0,45$ para F pequeno. É digno de nota que distribuições do tipo lei de potência (assintóticas) têm sido identificadas em diversos tipos de atividades de pesquisa científica, incluindo número de artigos por pesquisador[113], citações de artigos e periódicos [114-116], número de referências por artigo[115], número de citações por pesquisador[117], número de artigos por periódico[118] e frequência de uso de periódicos em bibliotecas[119].

Na fig. 5.2b, vemos também $p(F)$ em comparação com uma distribuição q -Weibull (veja apêndice A),

$$p(F) = aF^{r-1} \left[1 + (q-1) \frac{F^r}{b} \right]^{-1/(q-1)}, \quad (5.2)$$

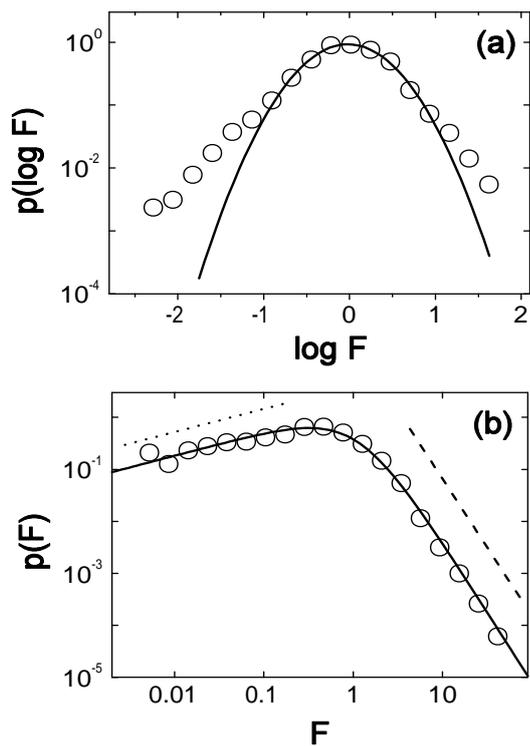


Figura 5.2: **(a)** Densidade de probabilidade $p(\log F)$, calculada para 5912 periódicos em 2004. A linha sólida representa uma distribuição Gaussiana. **(b)** Densidade de probabilidade $p(F)$ para os mesmos dados descritos em (a). Para comparação, mostramos as funções $p(F) \propto F^{-2,7}$ (linha tracejada) e $p(F) \propto F^{0,40}$ (linha pontilhada). A linha sólida representa a eq. (5.2), com $a = 1,45$, $b = 0,53$, $r = 1,45$ e $q = 1,46$.

com $r = 1,45$ e $q = 1,46$. Observe que a eq. (5.2) exibe comportamento assintótico $p(F) \sim F^{0,45}$ para pequenos valores de F e $p(F) \sim F^{-2,7}$ para grandes valores de F , reproduzindo assim o comportamento descrito na eq. (5.1). Repetimos esta análise (fig. 5.2) para cada um dos 13 anos considerados (1992-2004) e obtivemos resultados similares.

5.3 Distribuição das taxas de crescimento

Investigamos também propriedades estatísticas das taxas de crescimento

$$R(t) = \log \left[\frac{F(t+1)}{F(t)} \right], \quad (5.3)$$

com $F(t)$ e $F(t+1)$ sendo o fator de impacto nos anos t e $t+1$, respectivamente. Calculamos as taxas de crescimento $R(t)$ para cada um dos 2910 periódicos científicos num período de 13 anos entre 1992 e 2004. Dividimos então os periódicos em quatro grupos de acordo com seu valor inicial de F e, com base nestes grupos, calculamos a densidade de probabilidade condicional, $p(R|F)$, das taxas de crescimento do fator de impacto.

Na fig. 5.3a, vemos a densidade de probabilidade condicional $p(R|F)$ calculada para três grupos de periódicos. Observe que a distribuição exibe um comportamento aproximadamente exponencial, dado por

$$p(R|F) = \frac{1}{\sqrt{2}\sigma(F)} \exp \left(-\frac{\sqrt{2}|R - \mu|}{\sigma(F)} \right), \quad (5.4)$$

sendo μ e σ a média e o desvio padrão de R , respectivamente, calculados para o grupo em consideração. Observe que a forma da distribuição se mantém para os diferentes grupos considerados.

Para os mesmos grupos de periódicos, calculamos a distribuição das taxas de crescimento normalizadas

$$r(t) = \frac{R(t) - \mu}{\sigma}. \quad (5.5)$$

Na fig. 5.3b, vemos a distribuição condicional $p(r|F)$ em comparação com a função

$$p(r|F) = \frac{1}{\sqrt{2}} \exp(-\sqrt{2}|r|). \quad (5.6)$$

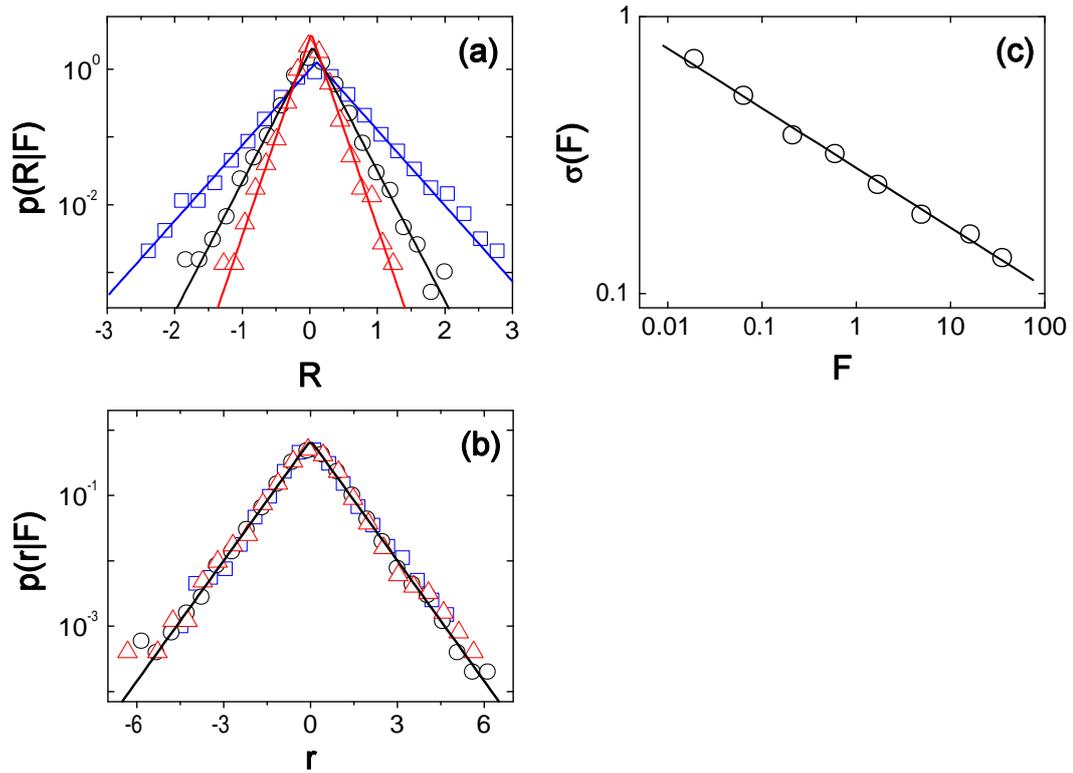


Figura 5.3: Dinâmica de crescimento das citações de periódicos científicos: pequenos e intermediários valores do fator de impacto. **(a)** Densidade de probabilidade condicional, $p(R|F)$, das taxas de crescimento R para três grupos de periódicos: 332 periódicos com $0,01 < F < 0,2$ (quadrados), 815 com $0,3 < F < 0,7$ (círculos) e 400 com $1,3 < F < 2$ (triângulos para cima). As linhas sólidas são dadas pela eq. (5.4) usando-se valores de μ e σ calculados diretamente a partir de R em cada grupo. **(b)** Densidade de probabilidade condicional, $p(r|F)$, das taxas de crescimento normalizadas r para os mesmos grupos de periódicos considerados em (a). Observe que as três curvas são bem descritas pela eq. 5.6. **(c)** Desvio padrão σ das taxas de crescimento R (calculado para cada grupo de periódicos) em função de F (valor médio do valor inicial de F). A linha sólida representa um ajuste linear, fornecendo um expoente $\beta \simeq 0,22$.

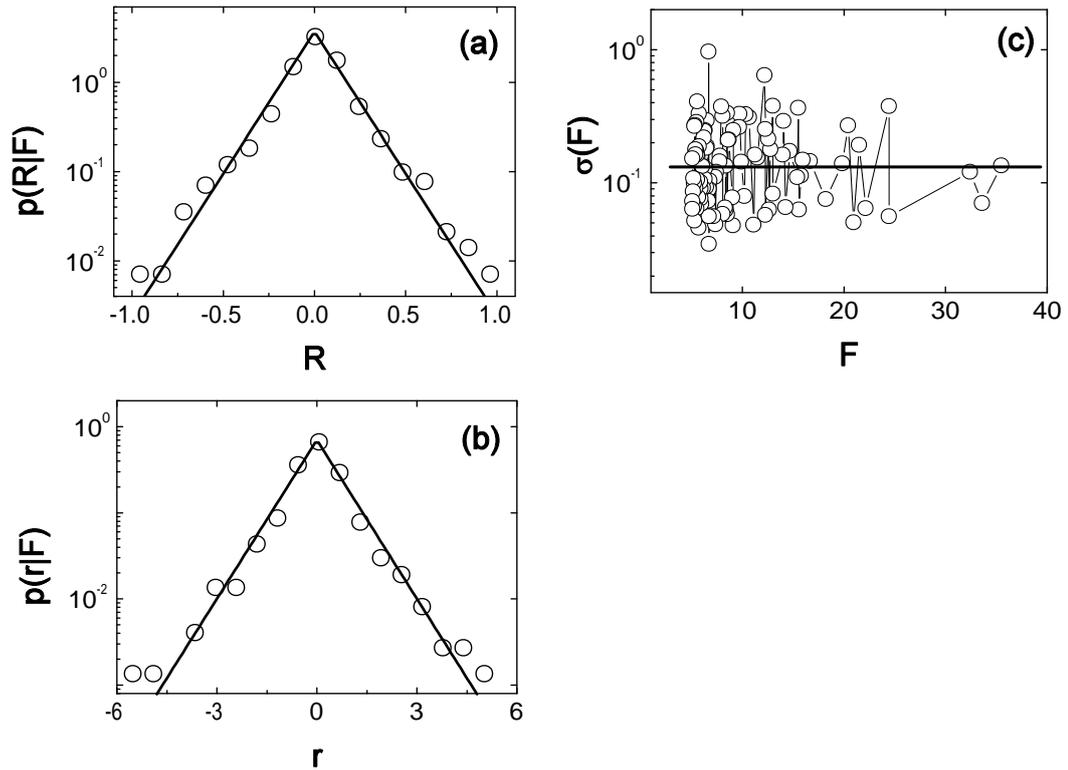


Figura 5.4: Dinâmica de crescimento das citações de periódicos científicos: os mais citados. **(a)** Densidade de probabilidade condicional $p(R|F)$ para um grupo de 100 periódicos, com $5,1 < F < 35,5$, no período 1992-2004. A linha sólida representa a eq. (5.4), com μ e σ calculados a partir dos dados. **(b)** Densidade de probabilidade condicional $p(r|F)$ calculada para os mesmos periódicos considerados em (a). A linha sólida representa a eq. (5.6). **(c)** Desvio padrão σ (calculado para cada um dos 100 periódicos considerados) das taxas R em função do fator de impacto inicial F . A linha horizontal indica que $\sigma(F) \sim F^{-\beta}$, com $\beta = 0$.

Note que as três curvas são bem descritas pela mesma função, dada pela eq. (5.6).

Na fig. 5.3a, vemos que as flutuações nas taxas de crescimento R diminuem com o aumento de F . Para investigar a forma desta dependência, calculamos o desvio padrão σ das taxas R em termos de F . Como vemos na fig. 5.3c,

$$\sigma(F) \propto F^{-\beta}, \quad (5.7)$$

com $\beta \simeq 0,22$.

Repetimos esta análise para o quarto grupo de periódicos: os mais citados (com maior valor de F). No que se refere à distribuição das taxas de crescimento, observamos na fig. 5.4a que $p(R|F)$ exhibe decaimento exponencial bem aproximado pela eq. (5.4). De modo similar, a distribuição das taxas de crescimento normalizadas, $p(r|F)$, se ajusta à função dada pela eq. (5.6), conforme ilustra a fig. 5.4b. Este comportamento é basicamente o mesmo observado nos outros três grupos de periódicos com pequenos e intermediários valores de F .

Por outro lado, observamos um comportamento distinto de $\sigma(F)$ para periódicos com grandes valores de F . Na fig. 5.4c, vemos o desvio padrão $\sigma(F)$ calculado para cada periódico dentro do grupo dos mais citados. Note que

$$\sigma(F) \sim F^{-\beta}, \quad (5.8)$$

com $\beta \simeq 0$ para grandes valores de F , indicando que nesta região não há uma significativa dependência de σ com F . Este comportamento não é observado na fig. 5.3c talvez porque o cálculo de σ para grupos de periódicos não seja suficientemente sensível para detectar esta mudança de comportamento.

Na próxima seção consideraremos características universais na dinâmica do fator de impacto.

5.4 Características universais

Os principais resultados obtidos nas seções anteriores foram: *i*) a distribuição $p(F)$ exhibe comportamento assintótico do tipo lei de potência, com expoente $\alpha \simeq 2,7$; *ii*) a distribuição das taxas de crescimento $p(R|F)$ exhibe decaimento exponencial, independentemente do valor de F ; *iii*) o desvio padrão $\sigma(F)$ decresce como uma lei de potência, com expoente $\beta \simeq 0,22$, passando a exhibir $\beta \simeq 0$ para grandes valores de F . Nas tabs.

5.1 e 5.2 vemos estes resultados em comparação com resultados descritos nos capítulos 3 e 4.

Tabela 5.1 - Dinâmica de crescimento do fator de impacto de periódicos científicos em comparação com a dinâmica do tamanho de universidades (capítulo 3).

	X	$p(\log X)$	$\sigma(X)$	β	$p(R X)$
Periódicos	citações	não-Gaussiana	$\propto X^{-\beta}$	0,22	exponencial
Universidades (EUA)	fundo $R\&D$	Gaussiana	$\propto X^{-\beta}$	0,25	exponencial
Universidades (EUA)	artigos	Gaussiana	$\propto X^{-\beta}$	0,25	exponencial
Universidades (EUA)	patentes	Gaussiana	$\propto X^{-\beta}$	0,25	exponencial

Tabela 5.2 - Dinâmica de crescimento do fator de impacto dos periódicos mais citados em comparação com a dinâmica das revistas mais lidas (capítulo 4) e das firmas que mais vendem (capítulo 3).

	X	$p(X)$	α	$\sigma(X)$	β	$p(R X)$
Periódicos	citações	$\sim X^{-\alpha}$	2,7	$\sim X^{-\beta}$	0	exponencial
Revistas	circulação	$\sim X^{-\alpha}$	2,5	$\sim X^{-\beta}$	0	exponencial
Firmas	vendas	$\sim X^{-\alpha}$	2,0	$\sim X^{-\beta}$	0	exponencial

Esses resultados indicam que a dinâmica de crescimento do fator de impacto de periódicos científicos exhibe padrões similares aos observados nas atividades econômicas e de pesquisa científica consideradas no capítulo 3 (compare as figs. 5.3, 3.1 e 3.2). Ademais, a dinâmica dos periódicos mais citados comporta-se como a dinâmica das revistas mais lidas e das firmas que mais vendem.

A seguir, sugeriremos um processo estocástico capaz de reproduzir aspectos do comportamento observado na dinâmica de crescimento do fator de impacto: a dependência de σ com F e a forma da distribuição $p(R|F)$. Naturalmente, o modelo também poderia se aplicar às outras organizações mencionadas nesta seção. Segue uma descrição do modelo.

5.5 Modelo estocástico II

Suponha que o fator de impacto $F(t)$ de cada periódico seja dado pelo processo estocástico multiplicativo e não linear

$$F(t+1) - F(t) = \lambda(t)[F(t)]^k, \quad (5.9)$$

com $k > 0$ e $F \geq F_m$, sendo F_m uma constante positiva e diferente de zero. Suponha também que $\lambda(t)$ seja dado por um processo multiplicativo do tipo

$$\lambda(t) = [b_0 + b_1\lambda(t-1)]\epsilon(t), \quad (5.10)$$

com $b_0 \ll 1$ e b_1 sendo constantes positivas. O termo $\epsilon(t)$ refere-se a um número aleatório cuja distribuição é gaussiana, com média zero e variância unitária, $\epsilon(t) = N(0, 1)$. O processo estocástico definido pelas eqs. (5.9) e (5.10) recupera o modelo de Gibrat (na ausência do corte F_m) quando $k = 1$ e $b_1 = 0$.

Verificamos que esta generalização do modelo de Gibrat é consistente com a equação $\sigma(F) \sim F^{-\beta}$, com $\beta > 0$, indicando que o desvio padrão das taxas de crescimento são dependentes do tamanho. Verificamos também que o valor médio das taxas de crescimento exibe um comportamento geral independente do tamanho, flutuando em torno de zero, principalmente para valores de F distantes da barreira refletora definida por F_m .

A dinâmica do modelo é controlada basicamente por dois parâmetros: k e b_1 . O efeito destes parâmetros nas previsões do modelo estão ilustrados na fig. 5.5, onde vemos resultados de simulações para diferentes valores de k e b_1 .

O parâmetro b_1 atua de modo análogo ao parâmetro a_1 do modelo proposto no capítulo anterior, proporcionando um efeito de memória. Este parâmetro é o principal responsável pela forma da distribuição das taxas de crescimento, gerando distribuições Gaussianas se $b_1 = 0$ e não-Gaussianas se $b_1 \neq 0$. Quanto maior o valor de $0 \leq b_1 \leq 1$, mais a distribuição se desvia da Gaussiana (fig. 5.5a). Por outro lado, a forma da distribuição das taxas de crescimento é praticamente independente de k (fig. 5.5b).

O parâmetro k reflete o fato de que as taxas de crescimento dependem do tamanho F . Se $k < 1$ as flutuações nas taxas de crescimento são amplificadas para F pequeno e amortecidas para F grande. Por outro lado, se $k > 1$ as flutuações são amortecidas para F pequeno e amplificadas para F grande. Deste modo, o parâmetro k governa

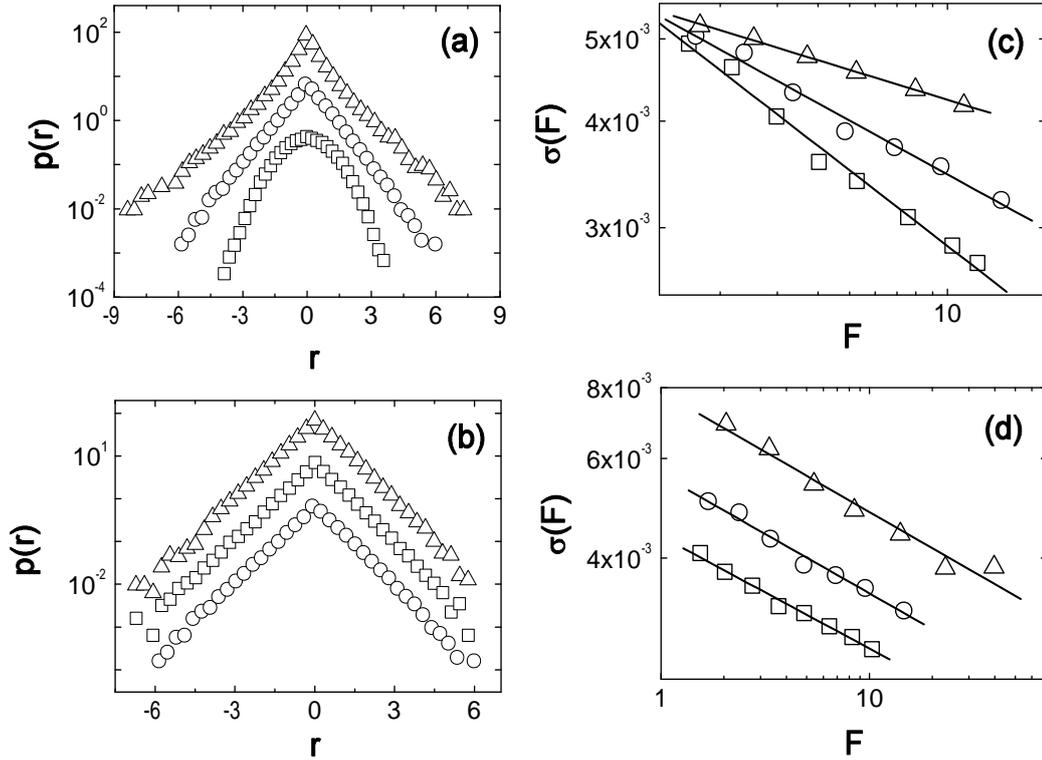


Figura 5.5: Realização típica do modelo definido pelas eqs. (5.9) e (5.10) para 1000 periódicos, todos com fator de impacto inicial $F_0 = 1$, com $F \geq F_m = 1$. Fixamos $b_0 = 0.01$ e variamos os valores de b_1 e k . Em cada caso consideramos 100 passos depois de um regime inicial. **(a)** Densidade de probabilidade $p(r)$, para $k = 1$ e $b_1 = 0$ (quadrados), $b_1 = 0,6$ (círculos) e $b_1 = 0,9$ (triângulos). **(b)** Densidade de probabilidade das taxas de crescimento normalizadas, $p(r)$, para $b_1 = 0.6$ e $k = 1$ (círculos), $k = 0,8$ (quadrados) e $k = 0,5$ (triângulos). Em (a) e (b) as curvas foram deslocadas na vertical para melhor visualização. **(c)** Desvio padrão das taxas de crescimento, $\sigma(F)$, para $b_1 = 0,6$ e $k = 0,7$ (quadrados), $k = 0,8$ (círculos) e $k = 0,9$ (triângulos). As linhas retas são ajustes lineares representando leis de potência com expoente $\beta \simeq 0,3$, $\beta \simeq 0,2$ e $\beta \simeq 0,1$, respectivamente. **(d)** Desvio padrão das taxas de crescimento, $\sigma(F)$, para $k = 0,8$ e $b_1 = 0$ (quadrados), $b_1 = 0,6$ (círculos) e $b_1 = 0,8$ (triângulos). As linhas retas representam leis de potência com expoente $\beta \simeq 0,2$.

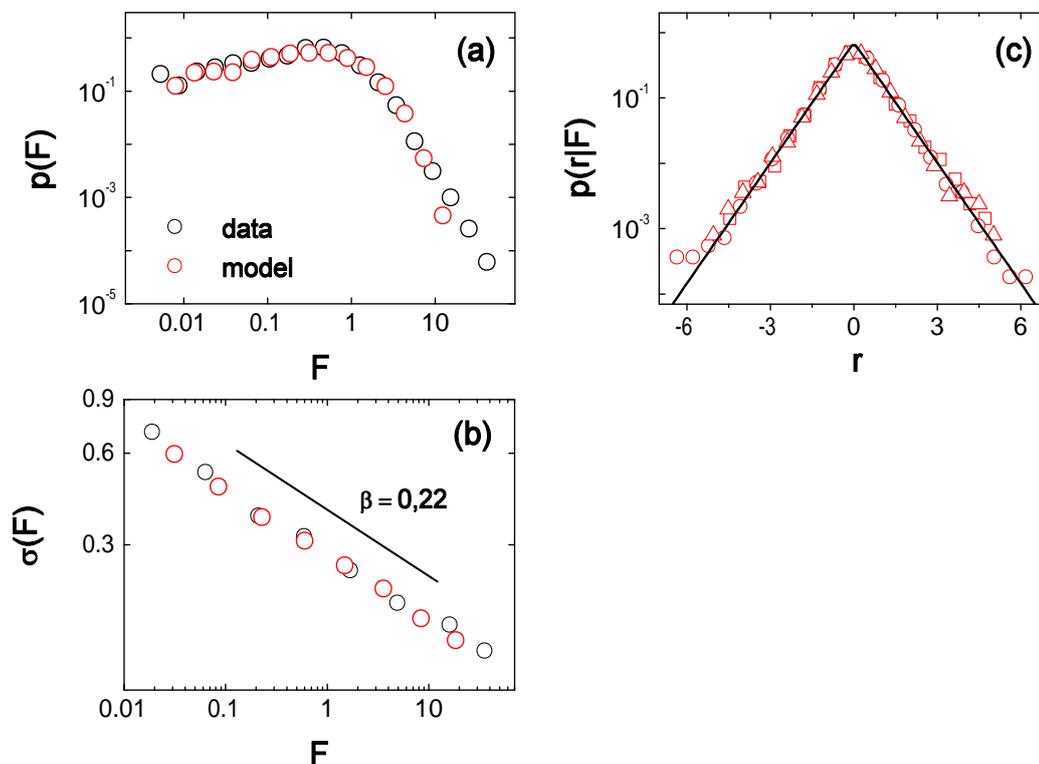


Figura 5.6: Realização típica do processo definido pelas eqs. (5.9) e (5.10) em comparação com resultados empíricos para o fator de impacto. Na simulação, usamos 5912 periódicos com $F_0 = 1,5$, $F_m = 0,005$ e $t = 13$ passos depois de descartar os passos iniciais. Fixamos $b_0 = 0,01$, $b_1 = 0,5$ e $k = 0,78$. **(a)** Densidade de probabilidade $p(F)$ para os 5912 periódicos na simulação (círculos preenchidos). A distribuição empírica (círculos) refere-se ao fator de impacto de 5912 periódicos no período 1992-2004 (fig. 5.2b). Note que este resultado vale para um certo número de passos do processo. À medida que o número de passos aumenta, a predição do modelo foge do comportamento observado. **(b)** Desvio padrão das taxas de crescimento $\sigma(F)$, calculados a partir de 2910 periódicos na simulação (círculos preenchidos). A curva empírica (círculos) refere-se ao fator de impacto de 2910 periódicos no período 1992-2004 (fig. 5.3c). O resultado da simulação em (b) foi convenientemente deslocado verticalmente para melhor comparação. **(c)** Densidade de probabilidade condicional, $p(r|F)$, para três grupos de periódicos na simulação (os mesmos grupos definidos na fig. 5.3a e 5.3b). As três curvas são bem descritas pela eq. 5.6.

a dependência do desvio padrão σ com F (fig. 5.5c). As simulações indicam que este parâmetro pode ser relacionado com o expoente β (veja eq. (5.8)) pela relação $\beta \simeq 1 - k$. Por sua vez, o expoente β não apresenta dependência significativa com b_1 (fig. 5.5d).

Na fig. 5.6, comparamos uma realização deste processo estocástico para um dado conjunto de parâmetros com os principais resultados empíricos apresentados neste capítulo. A partir do expoente empírico $\beta = 0,22$, podemos escolher $k \simeq 1 - \beta = 0,78$. Ademais, escolhemos $b_1 = 0,5$ de tal modo que a distribuição das taxas de crescimento seja aproximadamente exponencial. Uma realização típica do processo estocástico para esta escolha particular de parâmetros pode ser vista na fig. 5.6 em comparação com resultados empíricos. Observe que o modelo reproduz satisfatoriamente os principais aspectos do comportamento observado.

Capítulo 6

Atividades religiosas das Testemunhas de Jeová

Investigaremos aqui propriedades estatísticas de atividades religiosas das Testemunhas de Jeová em diferentes países[22]. Especificamente, calcularemos *i*) a distribuição do número de membros por país; *ii*) a distribuição das taxas anuais de crescimento; e *iii*) o desvio padrão das taxas de crescimento em termos do número de membros. Repetiremos esta análise para três medidas adicionais de atividade religiosa[22]. Por fim, consideraremos uma possível interpretação dos resultados usando o modelo II, proposto no capítulo anterior.

6.1 Introdução

Conforme já comentado (no capítulo 3), a dinâmica de crescimento de atividades econômicas e de atividades de pesquisa científica exibem leis empíricas surpreendentemente similares entre si. Neste capítulo, exploramos a possibilidade de que este comportamento universal se estenda a atividades religiosas. Especificamente, analisamos a dinâmica de crescimento das Testemunhas de Jeová.

As Testemunhas de Jeová são uma religião Cristã mundial composta de mais de 6,7 milhões de membros ativos, organizados em cerca de 100 mil congregações em 236 países[120]. Este grupo religioso é conhecido por suas crenças distintivas e pelas suas atividades de divulgação e ensino [121-124].

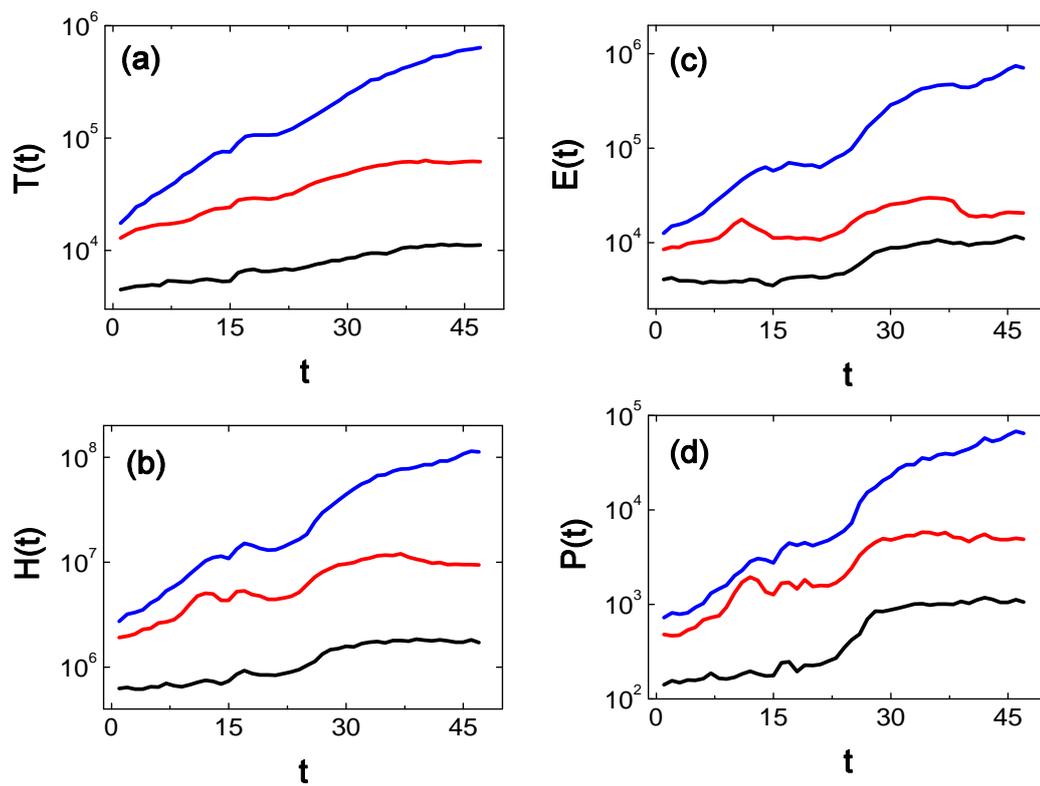


Figura 6.1: **(a)** Evolução temporal (anual) do número total de membros, $T(t)$, no período 1959-2005, nos seguintes países: Brasil (linha azul), Austrália (linha vermelha) e Jamaica (linha preta). **(b)** O mesmo que em (a) mas para o número total de horas, $H(t)$. **(c)** O mesmo que em (a) mas para a média de cursos bíblicos dirigidos, $E(t)$. **(d)** O mesmo que em (a) mas para a média de pioneiros, $P(t)$.

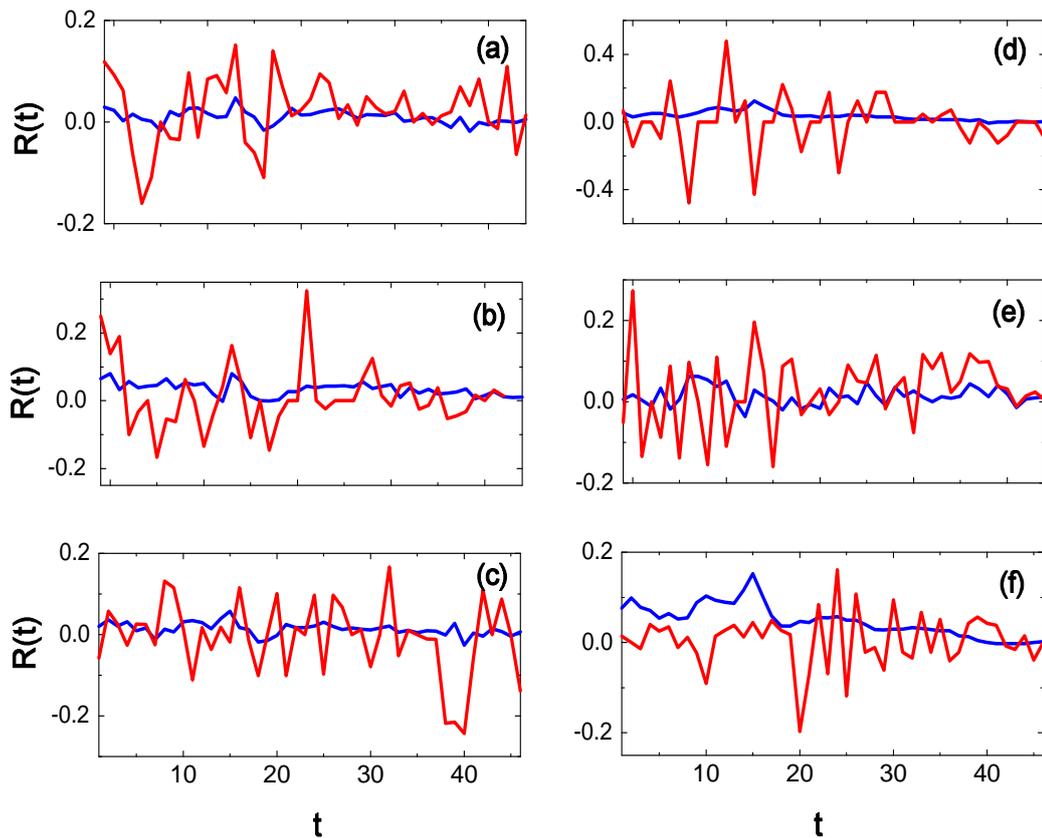


Figura 6.2: Evolução temporal das taxas de crescimento $R(t) = \log[T(t+1)/T(t)]$ no período 1959-2005 para os seguintes países: **(a)** Grã Bretanha (azul) e Ilhas Caimã (vermelho); **(b)** Brasil (azul) e Niue (vermelho); **(c)** Estados Unidos (azul) e Montserrat (vermelho); **(d)** Itália (azul) e Ilhas Malvinas (vermelho); **(e)** Zâmbia (azul) e Gâmbia (vermelho); **(f)** Japão (azul) e Névis (vermelho). Observe que as flutuações nas taxas de crescimento são maiores para países pequenos (com poucos membros).

O trabalho voluntário realizado pelas Testemunhas de Jeová inclui fazer visitas regulares às pessoas na sua localidade, usualmente de casa em casa e dirigir cursos bíblicos gratuitos com os interessados. Membros que se dedicam mais intensamente a esta obra de divulgação e ensino (em tempo parcial ou integral) são chamados pioneiros. Medidas desta atividade religiosa são meticulosamente relatadas ao redor do mundo e divulgadas (por país) em relatórios de serviço anuais[120]. Estas bases de dados têm sido consideradas fontes confiáveis - mas praticamente inexploradas - de informação sobre atividades religiosas[125].

De acordo com o relatório mundial de 2006[120], por exemplo, ao todo 6.741.444 membros dedicaram 1.333.966.199 horas ao serviço voluntário de divulgação e ensino, dirigiram em média 6.286.618 cursos bíblicos e, em média, 911.739 atuaram como pioneiros. A assistência à sua principal reunião anual foi de 16.675.113 de pessoas.

Neste estudo, investigamos propriedades estatísticas das seguintes medidas anuais: *i*) número total de membros, T ; *ii*) total de tempo gasto (em horas) na divulgação e ensino, H ; *iii*) média de cursos ou estudos dirigidos, E ; *iv*) média do número de pioneiros, P . Os dados analisados referem-se a 140 países no período de 47 anos, de 1959 a 2005. Vemos na fig. 6.1 a evolução temporal de T , H , E e P para três países no período considerado. Por sua vez, a fig. 6.2 mostra a evolução temporal das taxas de crescimento do número total de membros para alguns países selecionados no período 1959-2005. No que segue, passamos a descrever os resultados obtidos.

6.2 Distribuição do número de membros e de suas taxas de crescimento

Começamos calculando a distribuição do logaritmo do número de membros T , para 207 países em 2005. Na fig. 6.3a, vemos que $p(\log T)$ é consistente com uma distribuição Gaussiana,

$$p(\log T) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(\log T - \mu)^2}{2\sigma^2} \right], \quad (6.1)$$

sendo μ e σ a média e o desvio padrão de $\log T$, implicando que $p(T)$ é aproximadamente log-normal. De modo similar, a distribuição $p(\log T)$ calculada para outros anos também exhibe comportamento Gaussiano.

No que se refere à dinâmica de crescimento do número de membros das Testemunhas

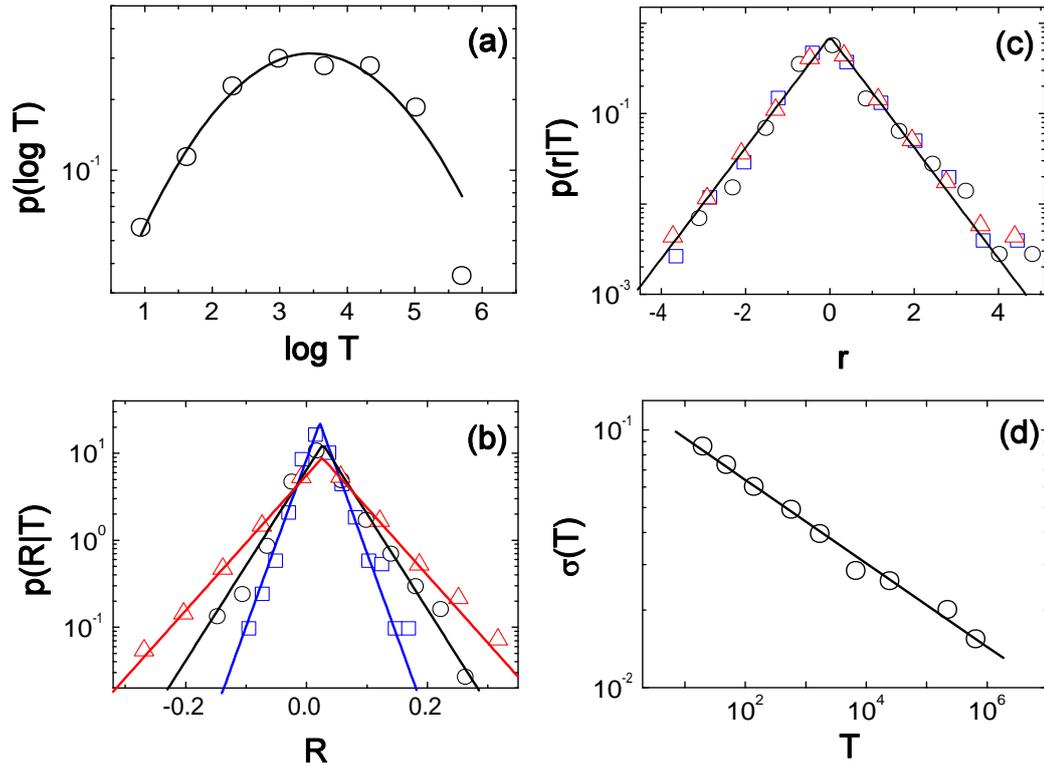


Figura 6.3: Dinâmica de crescimento do número de membros. (a) Densidade de probabilidade $p(\log T)$, sendo T o número de membros, para 207 países em 2005. A linha sólida representa uma distribuição Gaussiana, indicando que $p(T)$ é aproximadamente log-normal. (b) Densidade de probabilidade condicional, $p(R|T)$, calculada para três grupos de países: $6 < \langle T \rangle < 105$ (triângulos para cima), $105 < \langle T \rangle < 425$ (círculos) and $980 < \langle T \rangle < 2160$ (quadrados). As linhas sólidas são dadas pela eq. 6.3, com μ e σ calculados diretamente a partir dos dados. (c) Densidade de probabilidade condicional, $p(r|T)$, para os três grupos de países considerados em (b). Observe que as três curvas são bem descritas pela eq. 6.5. (d) Desvio padrão $\sigma(T)$ calculados para nove grupos de países. A linha sólida representa a eq. 6.8 com expoente $\beta = 0,16$.

de Jeová, calculamos a densidade de probabilidade condicional, $p(R|T)$, das taxas de crescimento

$$R(t) = \log \left[\frac{T(t+1)}{T(t)} \right]. \quad (6.2)$$

Para isso, consideramos três grupos de países selecionados de acordo com o valor médio de T no período 1959-2005. Na fig. 6.3b, vemos que as curvas empíricas $p(R|T)$ apresentam decaimento exponencial, sendo bem ajustadas pela distribuição

$$p(R|T) = \frac{1}{\sqrt{2}\sigma(T)} \exp \left(-\frac{\sqrt{2}|R - \mu|}{\sigma(T)} \right), \quad (6.3)$$

sendo μ e σ a média e o desvio padrão de R , respectivamente, dentro do grupo de países em questão. Note que a forma da distribuição se mantém praticamente inalterada. Entretanto, a distribuição apresenta diferentes larguras (desvio padrão) para cada um dos grupos de países.

A fig. 6.3c mostra a distribuição $p(r|T)$, das taxas de crescimento normalizadas

$$r(t) = \frac{R(t) - \mu}{\sigma}, \quad (6.4)$$

calculadas para os mesmos grupos de países considerados, em comparação com a equação

$$p(r|T) = \frac{1}{\sqrt{2}} \exp(-\sqrt{2}|r|). \quad (6.5)$$

Observe que as distribuições empíricas, calculadas para grupos distintos de países, são bem descritas por uma só curva.

Calculamos também o desvio padrão σ das taxas de crescimento R em termos de T . Conforme ilustra a fig. 6.3d, σ decresce com T seguindo uma lei de potência,

$$\sigma(T) \propto T^{-\beta}, \quad (6.6)$$

com $\beta \simeq 0,16$.

Na fig. 6.3a, vimos que $p(\log T)$ segue aproximadamente uma distribuição Gaussiana indicando que $p(T)$ é log-normal. Entretanto, observe que a distribuição parece se desviar deste comportamento para grandes valores de T . Para investigar mais detalhadamente o comportamento assintótico da distribuição do número de membros, calculamos a distribuição acumulada $p_c(T)$ para os países com maior número de membros em 2005. A fig. 6.4a sugere que $p_c(T)$ exibe comportamento assintótico do tipo

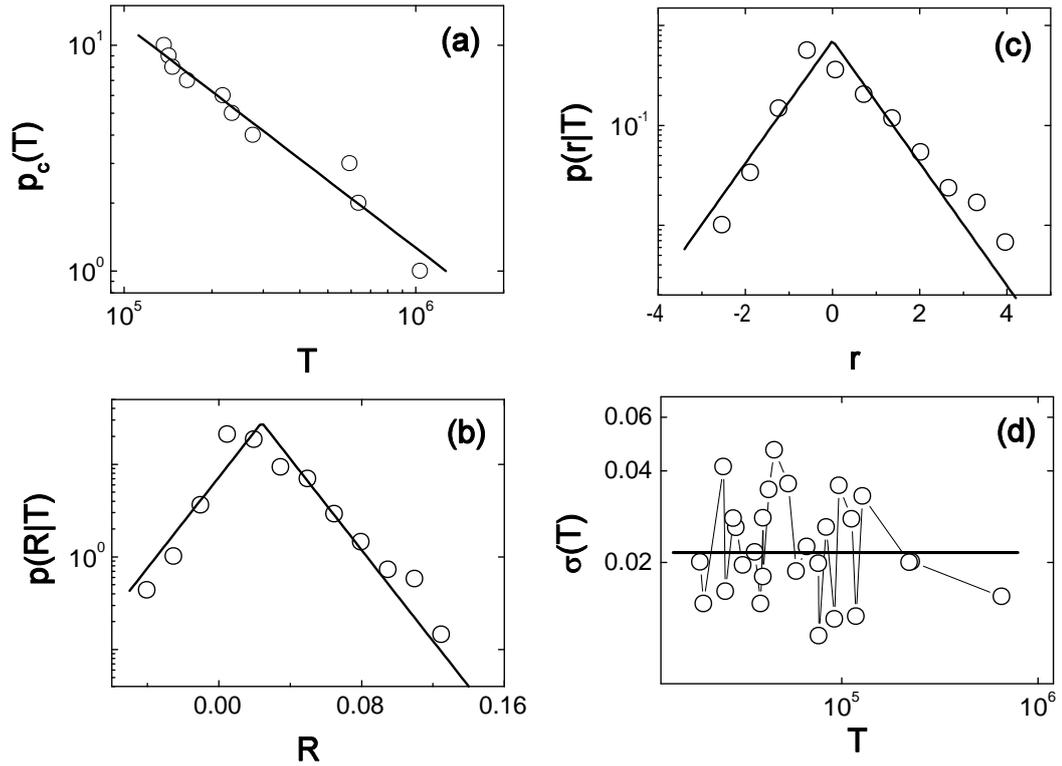


Figura 6.4: Dinâmica de crescimento do número de membros: países com muitos membros. **(a)** Distribuição acumulada, $p_c(T)$, para os 10 países com maior número de membros em 2005. A linha reta representa uma lei de potência, com expoente $\eta = 1,0$, indicando que $p(T) \sim T^{-\alpha}$, com $\alpha \simeq 2,0$. **(b)** Densidade de probabilidade condicional, $p(R|T)$, para países com $76.638 < \langle T \rangle < 655.607$ no período 1959-2005. A linha sólida representa a eq. (6.3), com μ e σ calculados diretamente a partir dos dados. **(c)** Densidade de probabilidade condicional, $p(r|T)$, para os mesmos dados considerados em (b). A linha sólida representa a eq. (6.5). **(d)** Desvio padrão $\sigma(T)$ calculado para os 27 países com maior valor médio de T no período considerado. A linha horizontal representa a função $\sigma(T) \sim T^{-\beta}$, com $\beta = 0$.

lei de potência,

$$p_c(T) \sim T^{-\eta}, \quad (6.7)$$

com $\eta \simeq 1, 0$. Isso indica que $p(T) \sim T^{-\alpha}$, com $\alpha \simeq 2, 0$.

A seguir, investigamos a densidade de probabilidade condicional, $p(R|T)$, para um quarto grupo de países: os países com maior valor médio de membros no período 1959-2005. Assim como se dá com os demais grupos de países, $p(R|T)$ exibe um decaimento exponencial, bem descrito pela eq. (6.3), conforme vemos na fig. 6.4b. A distribuição $p(r|T)$, para o mesmo grupo considerado, pode ser vista na fig. 6.4c, em comparação com a eq. (6.5).

Calculamos também $\sigma(T)$ para países com grandes valores de T . A fig. 6.4d sugere que $\sigma(T)$

$$\sigma(T) \sim T^{-\beta}, \quad (6.8)$$

com $\beta \simeq 0$ para grandes valores de T , indicando que nesta região σ não apresenta dependência significativa com T .

6.3 Tempo gasto, cursos dirigidos e pioneiros

Para testar se os resultados descritos na seção anterior valem para outras medidas de atividade religiosa, repetimos o mesmo tipo de análise para as demais variáveis consideradas. Verificamos que o análogo da fig. 6.3 vale para cada uma destas medidas.

Especificamente, *i*) as distribuições $p(\log H)$, $p(\log E)$ e $p(\log P)$ seguem aproximadamente um comportamento Gaussiano, indicando que as respectivas distribuições de tamanho apresentam comportamento log-normal; *ii*) as distribuições $p(R|H)$, $p(R|E)$ e $p(R|P)$ exibem decaimento exponencial, com larguras distintas, bem descritas por equações análogas à eq. (6.3); *iii*) as distribuições $p(r|H)$, $p(r|E)$ e $p(r|P)$ são bem descritas por uma só curva, conforme ilustram as figs. 6.5a, 6.5b e 6.5c; e *iv*) o desvio padrão σ das taxas de crescimento R decresce como uma lei de potência com o tamanho: $\sigma(H) \propto H^{-\beta}$ com $\beta \simeq 0, 16$; $\sigma(E) \propto E^{-\beta}$, com $\beta \simeq 0, 13$; e $\sigma(P) \propto P^{-\beta}$, com $\beta \simeq 0, 10$ (veja 6.5d).

Verificamos também que o análogo da fig. 6.4 vale para horas, cursos e pioneiros. Especificamente, *i*) as distribuições acumuladas $p_c(H)$, $p_c(E)$ e $p_c(P)$ exibem comportamento assintótico do tipo lei de potência, com expoentes $\eta \simeq 1$; *ii*) as distribuições $p(r|H)$, $p(r|E)$ e $p(r|P)$ são bem ajustadas pela mesma curva, conforme mostra a fig.

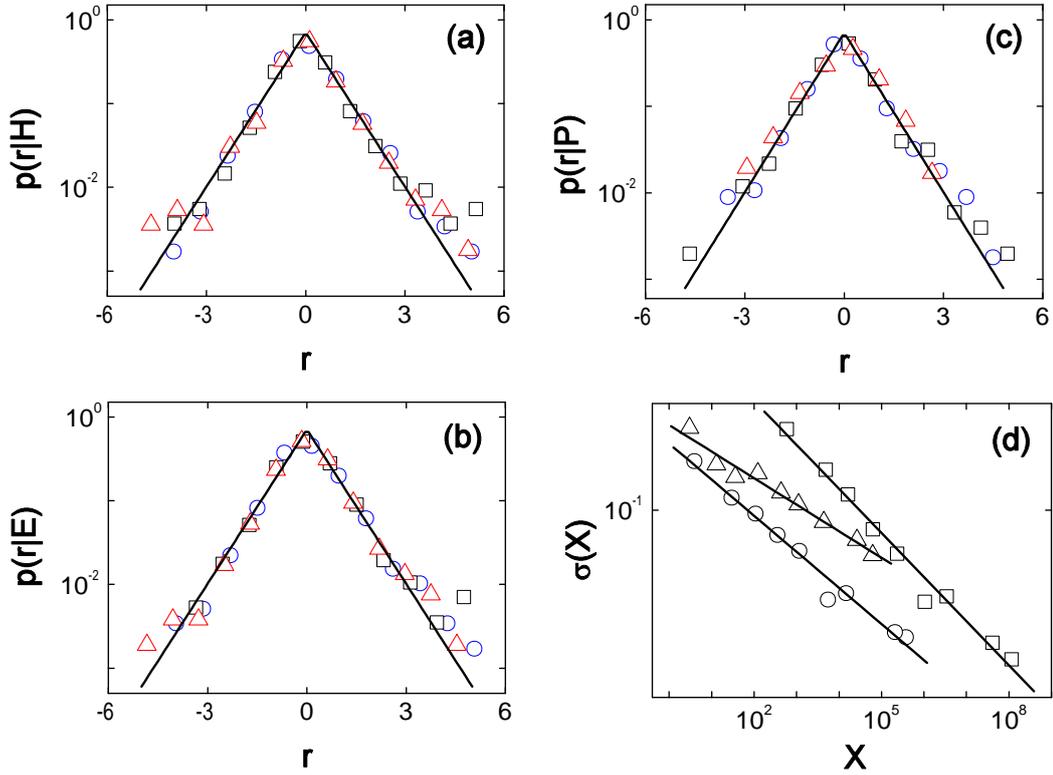


Figura 6.5: Dinâmica de crescimento do número de horas, cursos e pioneiros. **(a)** Distribuição condicional das taxas de crescimento do número de horas, $p(r|H)$, para os seguintes grupos de países: $2,30 \cdot 10^5 < \langle H \rangle < 3,66 \cdot 10^5$ (círculos); $3,87 \cdot 10^4 < \langle H \rangle < 7,78 \cdot 10^4$ (quadrados); e $5,70 \cdot 10^3 < \langle H \rangle < 2,27 \cdot 10^4$ (triângulos). **(b)** Distribuição condicional das taxas de crescimento do número de cursos, $p(r|E)$, para: $1,01 \cdot 10^3 < \langle E \rangle < 1,81 \cdot 10^3$ (círculos); $185 < \langle E \rangle < 379$ (quadrados); e $25 < \langle E \rangle < 121$ (triângulos). **(c)** Distribuição condicional das taxas de crescimento do número de pioneiros, $p(r|P)$, para: $114 < \langle P \rangle < 174$ (círculos); $25 < \langle P \rangle < 56$ (quadrados); e $3 < \langle P \rangle < 21$ (triângulos). **(d)** Desvio padrão $\sigma(H)$ (quadrados), $\sigma(E)$ (triângulos) e $\sigma(P)$ (círculos). As linhas sólidas representam leis de potência com expoentes $\beta \simeq 0,16$ para horas, $\beta \simeq 0,13$ para cursos e $\beta \simeq 0,10$ para pioneiros. A curva relacionada aos pioneiros foi deslocada verticalmente para melhor visualização.

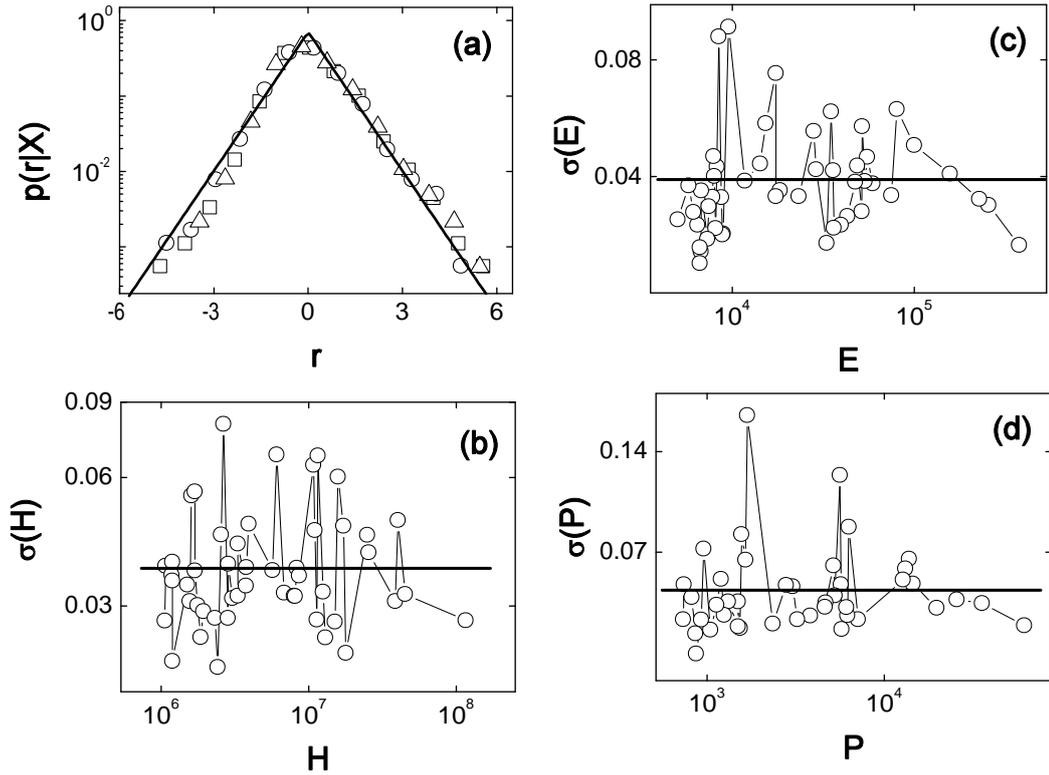


Figura 6.6: Dinâmica de crescimento do número de horas, estudos e pioneiros: países com grande atividade. **(a)** Densidade de probabilidade condicional, $p(r|H)$ (quadrados), para $1,06 \cdot 10^6 < \langle H \rangle < 1,16 \cdot 10^8$; $p(r|E)$ (círculos), para $5,02 \cdot 10^3 < \langle E \rangle < 3,77 \cdot 10^5$; e $p(r|P)$ (triângulos), para $7,39 \cdot 10^2 < \langle P \rangle < 6,19 \cdot 10^4$. A linha sólida representa a eq. (6.5). **(b)** Desvio padrão $\sigma(H)$ calculado para cada um dos países dentro do grupo considerado em (a). A linha horizontal representa a função $\sigma(H) \sim H^{-\beta}$, com $\beta = 0$. **(c)** O análogo dos resultados mostrados em (b), mas para $\sigma(E)$. **(d)** O análogo dos resultados mostrados em (b), mas para $\sigma(P)$.

6.6a; *iii*) o desvio padrão σ das taxas de crescimento R não apresenta uma dependência significativa com H , E e P nesta região, fornecendo $\beta \simeq 0$ (veja 6.6d). Estes resultados indicam que as mesmas leis gerais valem para diferentes medidas de atividades religiosas das Testemunhas de Jeová.

6.4 Aspectos universais

Os principais resultados descritos neste capítulo foram: *i*) a distribuição $p(\log T)$ é aproximadamente Gaussiana, indicando que a distribuição do número de membros $p(T)$ é bem ajustada por uma log-normal; *ii*) a distribuição das taxas de crescimento $p(R|T)$ exibe decaimento exponencial, independentemente do valor de T ; *iii*) o desvio padrão $\sigma(T)$ decresce como uma lei de potência, com expoente $\beta \simeq 0,16$; e *iv*) os resultados *i*, *ii* e *iii* valem para três outras medidas de atividade religiosa: tempo gasto ($\beta \simeq 0,16$), cursos dirigidos ($\beta \simeq 0,13$) e número de pioneiros ($\beta \simeq 0,10$). Estes resultados são resumidos na tab. 6.1.

Tabela 6.1 - Leis empíricas na dinâmica de crescimento de atividades religiosas das Testemunhas de Jeová.

	X	$p(\log X)$	$\sigma(X)$	β	$p(R X)$
Testemunhas de Jeová	membros	Gaussiana	$\propto X^{-\beta}$	0,16	exponencial
	horas	Gaussiana	$\propto X^{-\beta}$	0,16	exponencial
	cursos	Gaussiana	$\propto X^{-\beta}$	0,13	exponencial
	pioneiros	Gaussiana	$\propto X^{-\beta}$	0,10	exponencial

Na tab. 6.2, comparamos a dinâmica de atividades religiosas sem fins lucrativos com a dinâmica de atividades econômicas. Observe que as mesmas leis gerais valem para estes diferentes aspectos da atividade humana. Note também que a similaridade entre atividades religiosas e econômicas é quantitativa, incluindo o valor do expoente β .

Tabela 6.2 - Comparação entre a dinâmica de crescimento de atividades religiosas voluntárias e de atividades econômicas (capítulo 3).

	X	$p(\log X)$	$\sigma(X)$	β	$p(R X)$
Religião	membros	Gaussiana	$\propto X^{-\beta}$	0, 16	exponencial
Religião	horas	Gaussiana	$\propto X^{-\beta}$	0, 16	exponencial
Países	PIB	Gaussiana	$\propto X^{-\beta}$	0, 15	exponencial
Firmas	vendas	Gaussiana	$\propto X^{-\beta}$	0, 15	exponencial
Firmas	funcionários	Gaussiana	$\propto X^{-\beta}$	0, 16	exponencial

No que se refere ao grupo dos países com maior número de membros, verificamos que: *i*) a distribuição $p(T)$ apresenta comportamento assintótico do tipo lei de potência, com expoente $\alpha \simeq 2,0$; *ii*) a distribuição das taxas de crescimento $p(R|F)$ continua exibindo decaimento exponencial; e *iii*) o desvio padrão $\sigma(F)$ exibe um expoente $\beta \simeq 0$.

Estes resultados são resumidos na tab. 6.3, em comparação com resultados análogos observados na circulação de revistas, citações de periódicos e tamanho de firmas. Mais uma vez, observe as similaridades quantitativas entre atividades religiosas e econômicas.

Tabela 6.3 - Dinâmica de crescimento dos países com maior número de membros, dos periódicos mais citados (capítulo 5), das revistas mais lidas (capítulo 4) e das firmas que mais vendem (capítulo 3).

	X	$p(X)$	α	$\sigma(X)$	β	$p(R X)$
Religião	membros	$\sim X^{-\alpha}$	2,0	$\sim X^{-\beta}$	0	exponencial
Periódicos	citações	$\sim X^{-\alpha}$	2,7	$\sim X^{-\beta}$	0	exponencial
Revistas	circulação	$\sim X^{-\alpha}$	2,5	$\sim X^{-\beta}$	0	exponencial
Firmas	vendas	$\sim X^{-\alpha}$	2,0	$\sim X^{-\beta}$	0	exponencial

No que segue, compararemos alguns resultados apresentados neste capítulo com as predições do modelo II, proposto no capítulo anterior.

6.5 Comparação com o modelo estocástico II

Suponha que o número total de membros das Testemunhas de Jeová num dado país, $T(t)$, seja dado pelo processo estocástico descrito no capítulo anterior (modelo II). Neste caso,

$$T(t+1) - T(t) = \lambda(t)[T(t)]^k, \quad (6.9)$$

com $k > 0$ e $T \geq T_m$, sendo T_m o número mínimo de membros num dado país. Por sua vez, o ruído $\lambda(t)$ é dado pela equação

$$\lambda(t) = [c_0 + c_1\lambda(t-1)]\epsilon(t), \quad (6.10)$$

com $c_0 \ll 1$ e c_1 sendo constantes positivas. O número aleatório $\epsilon(t)$ segue uma distribuição Gaussiana com média zero e variância unitária.

A seguir, comparamos a evolução temporal das taxas de crescimento geradas pelo processo estocástico definido nas eqs. (6.9) e (6.10), para um conjunto particular de parâmetros, com a evolução das taxas de crescimento empíricas de alguns países selecionados. Na fig. 6.7, vemos que as predições do modelo seguem um padrão parecido com o comportamento empírico observado. Observe que, em ambos os casos as flutuações nas taxas de crescimento são menores para países com maior número de membros.

Além disso, repetimos o mesmo tipo de análise empregada nas figs. 6.3b, 6.3c e 6.3d, para o estudo das taxas de crescimento geradas pelo modelo. Especificamente, calculamos as distribuições $p(R|T)$ para três grupos distintos de países, selecionados de acordo com seu número médio de membros no período da simulação. Verificamos que $p(R|T)$ apresenta decaimento exponencial, bem aproximado pela eq. (6.3), conforme ilustra a fig. 6.8a. Além disso, para estes mesmos grupos de países, calculamos a distribuição das taxas de crescimento normalizadas, $p(r|T)$. Na fig. 6.8b vemos que as distribuições podem ser bem descritas por uma só curva, dada pela eq. (6.5). Por fim, dividimos os países considerados na simulação em oito grupos e calculamos o desvio padrão σ dentro de cada grupo em termos de T (fig. 6.8c). Verificamos que $\sigma(T) \propto T^{-\beta}$, com $\beta \simeq 0,16$. Portanto, vemos que o modelo reproduz satisfatoriamente os principais aspectos do comportamento observado.

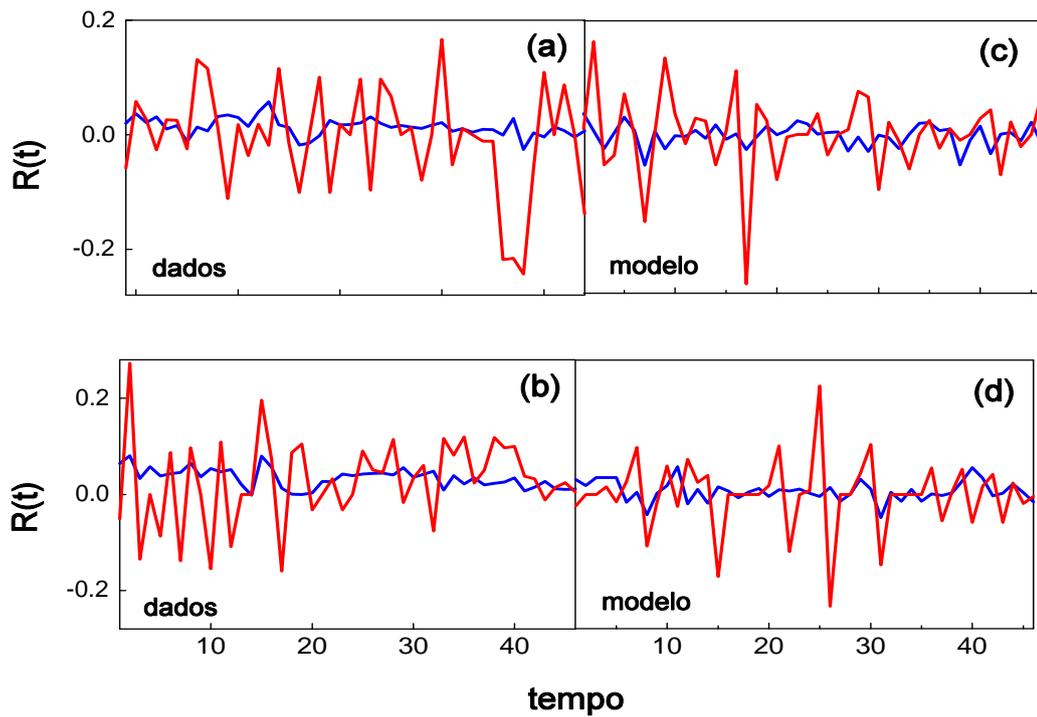


Figura 6.7: Evolução temporal das taxas de crescimento empíricas em comparação com as previstas pelo modelo II. Taxas de crescimento empíricas $R(t) = \log[T(t+1)/T(t)]$, no período 1959-2005, para: **(a)** Estados Unidos (círculos) e Montserrat (quadrados); **(b)** Brasil (círculos) e Gâmbia (quadrados). **(c)** e **(d)** Taxas de crescimento típicas geradas pelo processo definido nas eqs. (6.9) e (6.10), com $c_0 = 0,18$, $c_1 = 0,50$ and $k = 0,85$. Definimos o número mínimo de membros num país como sendo $T_m = 1$ e utilizamos o número empírico de membros no ano 1959 com sendo o número inicial de membros na simulação. Cada passo de tempo do modelo corresponde a um ano.

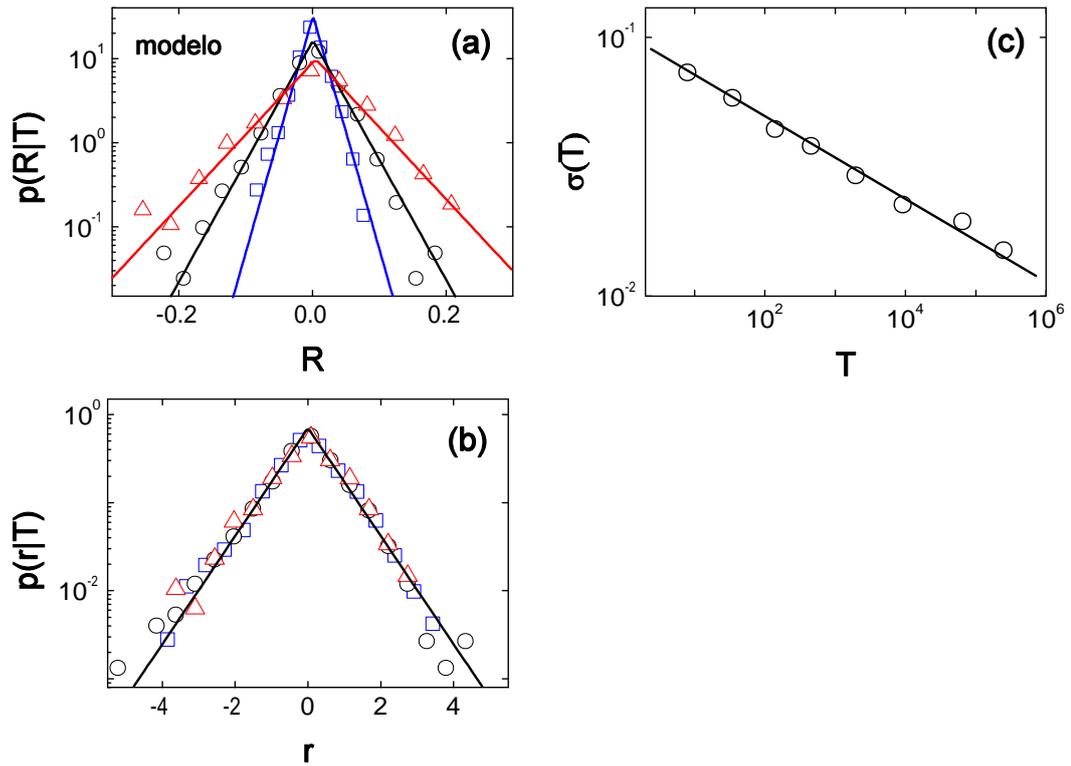


Figura 6.8: Simulação da dinâmica de crescimento do número de membros. Realização típica do processo estocástico definido pelas eqs. (6.9) and (6.10). Na simulação, consideramos 140 países com número inicial de membros obtidos a partir dos dados empíricos (relativos ao ano 1959) e os mesmos parâmetros utilizados na simulação da fig. 6.7. **(a)** Densidade de probabilidade condicional, $p(R|T)$, calculada para três grupos de países: com pequeno (quadrados), médio (círculos) e grande (triângulos) número de membros. As linhas sólidas são dadas pela eq. 6.3, com μ e σ calculados diretamente a partir dos dados. **(b)** Densidade de probabilidade condicional, $p(r|T)$, para os três grupos de países considerados em (a), em comparação com a eq. 6.5 (linha sólida). **(c)** Desvio padrão $\sigma(T)$ calculados para oito grupos de países. A linha sólida representa a eq. 6.8 com expoente $\beta = 0,16$.

Capítulo 7

Sinal geomagnético: o índice Dst

Investigaremos aqui a dinâmica de um sinal que mede a intensidade do campo magnético da Terra na região do equador - o índice Dst[23]. Especificamente, *i*) calcularemos a distribuição das flutuações no sinal em diferentes escalas de tempo; *ii*) analisaremos as correlações temporais utilizando o método *detrended fluctuations analysis* (DFA); *iii*) investigaremos propriedades multifractais do sinal utilizando o método *multifractal detrended fluctuations analysis* (mf-DFA). Ademais, compararemos a dinâmica deste sinal geomagnético com a dinâmica de um sinal fisiológico.

7.1 Introdução

O ambiente que cerca a Terra, constituído de partículas e campos controlados pelo campo magnético terrestre (a magnetosfera), não é estático e isolado no espaço. A magnetosfera se encontra imersa em uma região de plasma que flui constantemente do Sol, chamada de vento solar. Eventos eruptivos solares liberam grande quantidade de energia e partículas para o espaço. Neste respeito, a magnetosfera é essencial para a vida na Terra, atuando como uma barreira protetora. Entretanto, ao interagir com a magnetosfera, o vento solar provoca efeitos tais como a penetração de partículas, indução de correntes e alterações do campo magnético terrestre.

As tempestades geomagnéticas são um dos efeitos mais marcantes da interação vento solar-magnetosfera. Elas relacionam-se com a formação de correntes na magnetosfera. Em particular, a chamada corrente anelar (ou corrente de anel) produz um campo magnético de direção oposta ao campo magnético intrínseco da Terra na região

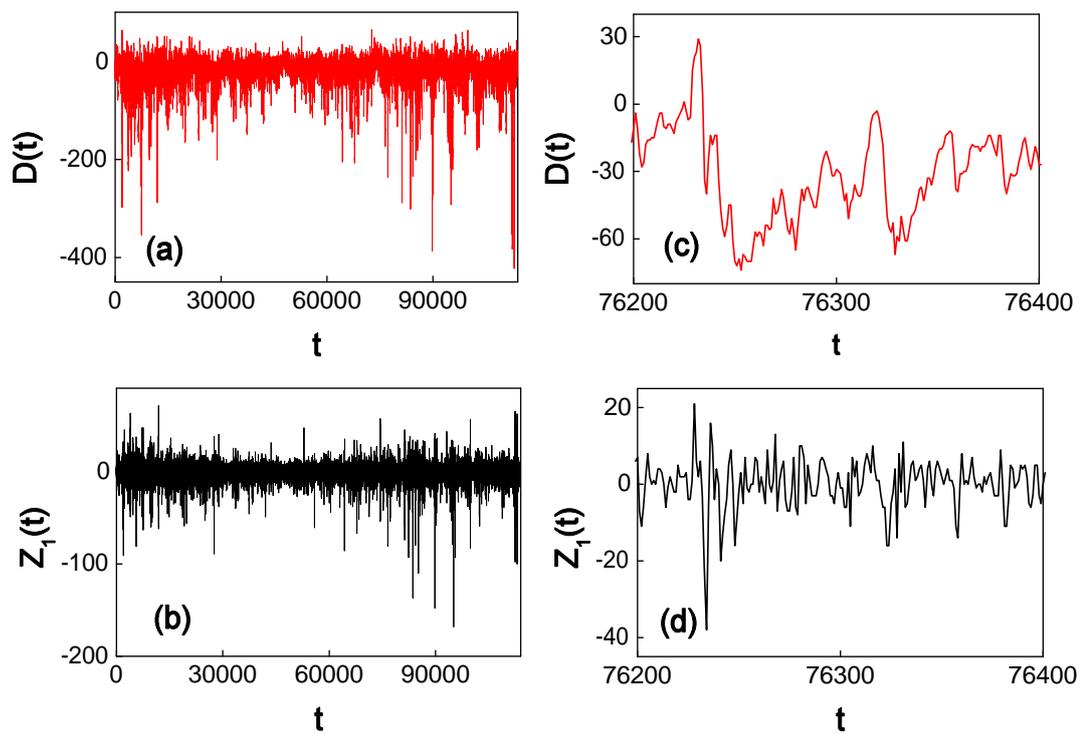


Figura 7.1: **(a)** Evolução temporal do índice Dst no período 1991-2003, sendo $D(t)$ o valor do índice no tempo t (em horas); **(b)** Incrementos consecutivos, $Z_1(t) = D(t + 1) - D(t)$, obtidos da série em (a); **(c)** Evolução temporal do índice Dst num período de 200 horas; **(d)** Incrementos consecutivos obtidos da série em (c).

do equador e, conseqüentemente, um decréscimo na componente horizontal do campo magnético da Terra. Este decréscimo, usualmente com duração de algumas horas, caracteriza a fase principal de uma tempestade geomagnética. Intensas tempestades geomagnéticas podem afetar sistemas tecnológicos tais como redes elétricas, oleodutos, transmissão de ondas de rádio de alta freqüência, sistemas de navegação, entre outros[126, 127]. Há evidências de que tempestades geomagnéticas também afetam sistemas biológicos, inclusive o sistema cardio-vascular humano[128, 129].

Variações da corrente anelar equatorial são monitoradas pelo índice Dst (*Disturbance storm time*), proposto no fim dos anos 1960s. O índice Dst é definido como o campo perturbado axialmente simétrico no equador do dipolo na superfície da Terra [130-132]. Expresso em nanoteslas, o Dst baseia-se no valor médio da componente horizontal do campo magnético da Terra medido de hora em hora em quatro observatórios geomagnéticos estrategicamente localizados. Este índice reflete variações na intensidade da parte simétrica da corrente anelar que circunda a Terra em altitudes no intervalo de 3-8 raios terrestres. O Dst é considerado uma estimativa global razoável de variações na componente horizontal do campo magnético perto do equador.

Nesse capítulo, analisamos propriedades estatísticas do índice Dst no período de 1991 a 2003[133]. Estes dados estão ilustrados na fig. 7.1, onde vemos a série temporal do índice Dst e de seus incrementos consecutivos.

7.2 Distribuição das flutuações no índice Dst

Iniciamos investigando as flutuações no índice Dst em diferentes escalas de tempo. Para isso, definimos os incrementos

$$Z_n(t) = D(t + n) - D(t), \quad (7.1)$$

sendo n a escala de tempo (em horas) e $D(t)$ o valor do índice Dst no tempo t . Para melhor comparação entre as distribuições dos incrementos Z_n , para diferentes valores de n , usamos a variável normalizada

$$z_n(t) = \frac{Z_n(t)}{\sigma}, \quad (7.2)$$

sendo σ o desvio padrão de Z_n .

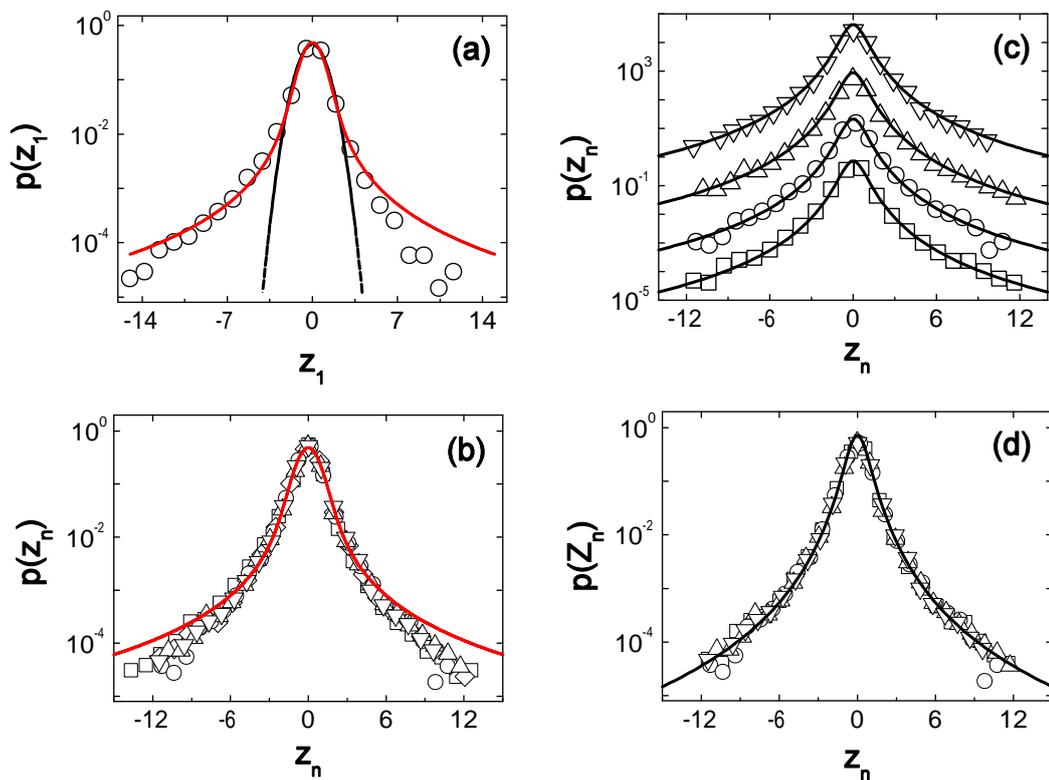


Figura 7.2: Distribuição das flutuações no índice Dst. **(a)** Densidade de probabilidade $p(z_1)$ calculada no período 1991-2003. A linha sólida representa a distribuição de Lévy, dada pela eq. (7.3), com $\psi = 1,7$. A linha tracejada representa a eq. (7.3), com $\psi = 2,0$ (Gaussiana). **(b)** Densidade de probabilidade $p(z_n)$ para diferentes escalas de tempo: $n = 36$ (quadrados), $n = 120$ (losangos), $n = 360$ (círculos), $n = 720$ (triângulos para cima) e $n = 1080$ (triângulos para baixo). A linha sólida representa uma distribuição de Lévy, com $\psi = 1,7$. **(c)** Densidade de probabilidade $p(z_n)$, para $n = 96$ (quadrados), $n = 360$ (círculos), $n = 720$ (triângulos para cima) e $n = 1080$ (triângulos para baixo). As linhas sólidas referem-se à distribuição q -Gaussiana, dada pela eq. (7.4), com $q = 1,5$. As curvas foram deslocadas verticalmente para melhor visualização. **(d)** As mesmas distribuições mostradas em (c), sem o deslocamento vertical, em comparação com a distribuição q -Gaussiana com $q = 1,5$.

Na fig. 7.2a, vemos a distribuição $p(z_1)$, calculada para o inteiro período 1991-2003, em comparação com uma distribuição simétrica de Lévy,

$$L(z) = \frac{1}{\pi} \int_0^\infty \exp(-\rho k^\psi) \cos(kz) dk, \quad (7.3)$$

com $\rho > 0$ (veja o apêndice A). Observe que a parte central de $p(z_1)$ pode ser aproximada por uma distribuição de Lévy, com $\psi = 1,7$. Entretanto, $p(z_1)$ desvia-se da distribuição de Lévy para grandes valores de z_1 . Ademais, calculamos a distribuição $p(z_n)$ para diferentes valores de n . Como pode ser visto na fig. 7.2b, a forma da distribuição se mantém praticamente inalterada para diferentes escalas de tempo, indicando a presença de invariância de escala na distribuição das flutuações na série Dst.

Comparamos também as distribuições empíricas das flutuações no índice Dst, para diferentes escalas de tempo, com a distribuição q -Gaussiana normalizada (veja o apêndice A),

$$p(z) = a \left[1 + (q-1) \frac{z^2}{b} \right]^{-1/(q-1)} \quad (7.4)$$

com $b = 5 - 3q$ e $q = 1,5$. Nas figs. 7.2c e 7.2d, vemos que a q -Gaussiana fornece uma boa descrição das principais características da distribuição empírica, incluindo o comportamento da parte central e o comportamento das caudas. Note que $p(z) \sim z^{-2/(q-1)} = z^{-4}$ para grandes valores de z , indicando que a distribuição das flutuações no índice Dst exibem decaimento do tipo lei de potência nas caudas, com expoente $\simeq 4$.

7.3 Correlações de longo alcance

A seguir, investigamos a presença de correlações de longo alcance na série $D(t)$ e na série dos seus incrementos consecutivos $Z_1(t)$. Em particular, a série $Z_1(t)$ pode ser decomposta em duas sub-séries: a série das magnitudes e a série dos sinais. Mais especificamente,

$$Z_1(t) = |Z_1(t)| \text{sign}[Z_1(t)], \quad (7.5)$$

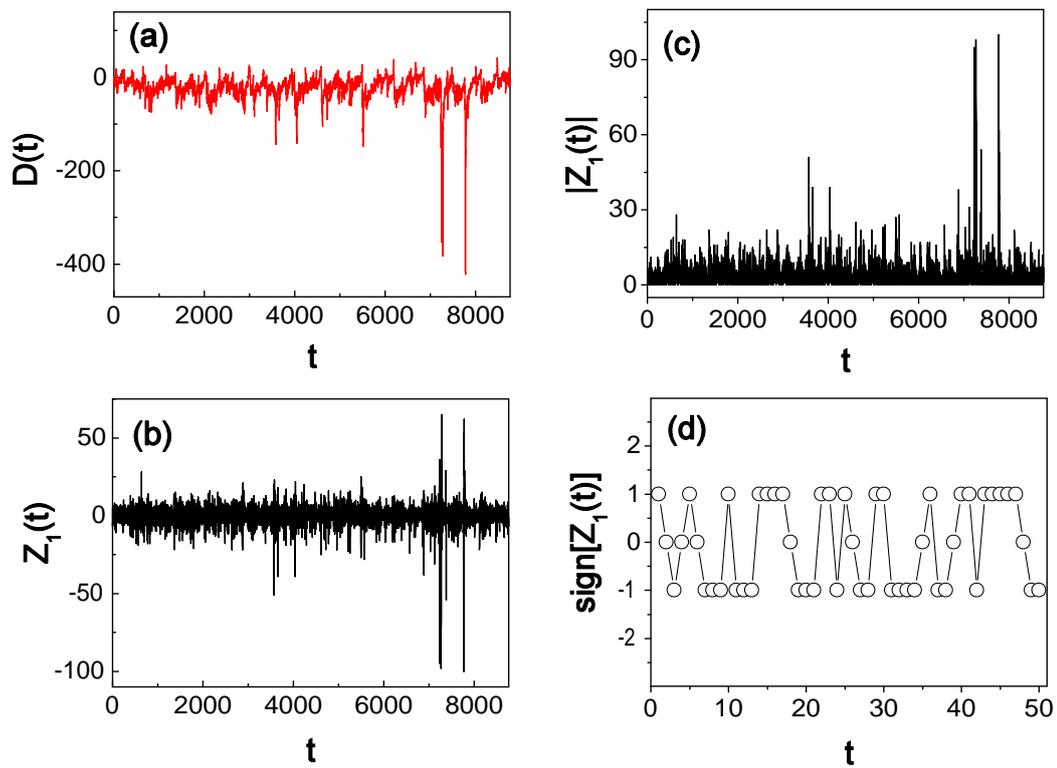


Figura 7.3: **(a)** Evolução temporal do índice Dst, $D(t)$, para o ano 2003. **(b)** Série dos incrementos consecutivos, $Z_1(t)$, obtida a partir da série em (a). **(c)** Série das magnitudes dos incrementos, $|Z_1(t)|$, obtida de (b). **(d)** Série dos sinais, $\text{sign}[Z_1(t)]$, para os 50 primeiros elementos em (b). A função sinal assume os valores -1 , 0 e 1 se o incremento for negativo, nulo ou positivo, respectivamente.

sendo $|Z_1(t)|$ o valor absoluto do incremento e

$$\text{sign}[Z_1(t)] = \begin{cases} 1, & Z_1(t) > 0 \\ 0, & Z_1(t) = 0 \\ -1, & Z_1(t) < 0 \end{cases} . \quad (7.6)$$

Na fig. 7.3, vemos as séries $D(t)$, $Z_1(t)$, $|Z_1(t)|$ e $\text{sign}[Z_1(t)]$ para o ano 2003.

Aplicamos o método DFA[28, 29], descrito no capítulo 2, para investigar a presença de correlações temporais nas séries ilustradas nas figs. 7.3a, 7.3c e 7.3d. Em cada caso, calculamos a função de flutuação $F(s)$, obtida a partir da versão integrada da série em questão. Vemos na fig. 7.4 que $F(s)$ cresce com a escala de tempo s como uma lei de potência,

$$F(s) \sim s^{h+1}, \quad (7.7)$$

com expoentes $h \simeq 1.0$ para a série do índice Dst (fig. 7.4a), $h \simeq 0,8$ para a série das magnitudes (fig. 7.4b) e $h \simeq 0,4$ para a série dos sinais (fig. 7.4c). Estes expoentes foram obtidos no intervalo $6 \leq s \leq 1080$ horas. Nestas figuras, também podem ser vistas as funções $F(s)$ calculadas para versões embaralhadas das séries consideradas. Como esperado, nestes casos $h \simeq 0,5$ indicando ausência de correlações de longo alcance.

Resultados análogos aos ilustrados na fig. 7.4 valem para cada um dos anos no período considerado (1991-2003). A evolução temporal dos expoentes h neste período pode ser vista na fig. 7.4d. Observe que $\langle h \rangle \simeq 1,0$ para a série Dst, $\langle h \rangle \simeq 0,75$ para a série das magnitudes e $\langle h \rangle \simeq 0,4$ para a série dos sinais.

No caso da série $D(t)$, investigamos também o comportamento de $F(s)$ no intervalo $6 \leq s \leq 96$. Nesta região, $h \simeq 1,3$, em contraste com o valor $h \simeq 1,0$ obtido no intervalo $6 \leq s \leq 1080$. O expoente $h > 1$ válido para pequenas escalas de tempo é consistente com resultados descritos na ref. [134], que sugerem que a dinâmica do índice Dst segue aproximadamente um ruído Browniano ($h \simeq 1,5$). Entretanto, deve-se ressaltar que os resultados apresentados nas figs. 7.4a e 7.4d sugerem que a dinâmica do índice Dst segue aproximadamente um ruído $1/f$ no intervalo considerado.

Os resultados ilustrados na fig. 7.4 indicam a presença de correlações de longo alcance na dinâmica do índice Dst. Especificamente, o resultado $h \simeq 1,0$ para a série $D(t)$ corresponde a um expoente $h \simeq 0$ para a série $Z_1(t)$. Isso indica que a série dos incrementos consecutivos é fortemente anti-correlacionada (anti-persistente). Ademais, a série das magnitudes exhibe $h > 0,5$, indicando comportamento persistente, enquanto que a série dos sinais exhibe $h < 0,5$, indicando anti-persistência. Estes resultados

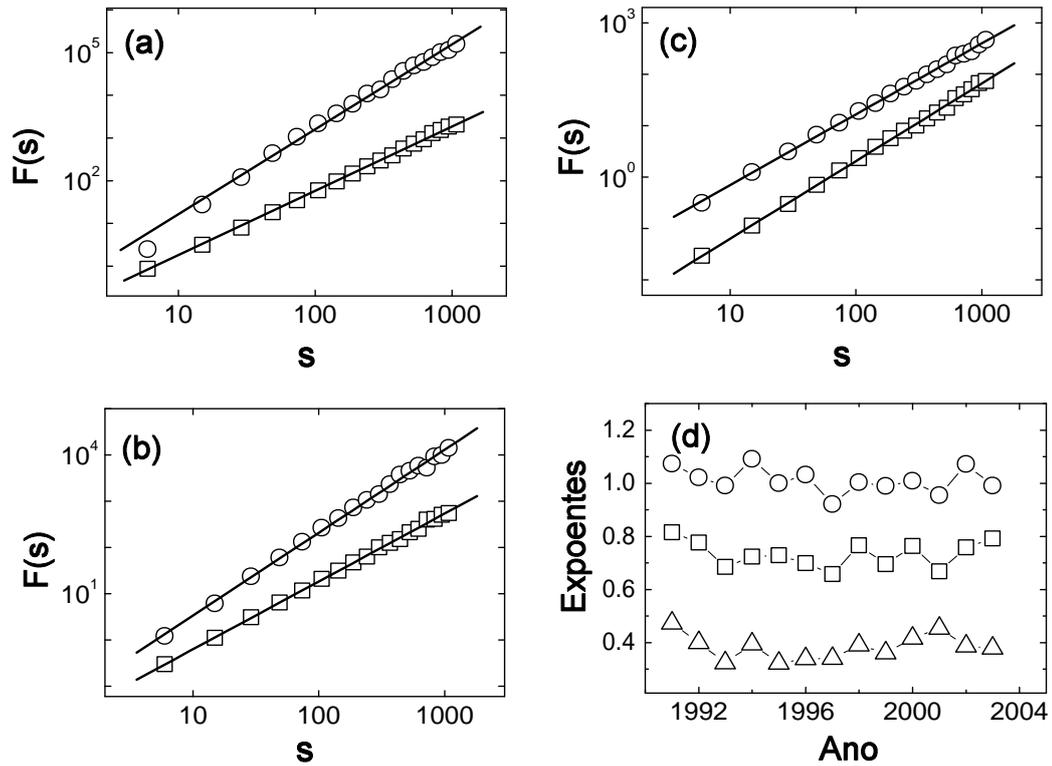


Figura 7.4: Correlações de longo alcance no índice Dst. $F(s)$ refere-se à função de flutuação obtida pelo uso do método DFA de ordem 2 nas respectivas séries integradas. **(a)** $F(s)$ para a série $D(t)$ em 2003 (círculos) e para sua versão embaralhada (quadrados). As linhas sólidas representam leis de potência, dadas pela eq. (7.7), com $h = 1,0$ para a série original e $h = 0,5$ para sua versão embaralhada. **(b)** O análogo de (a) para a série das magnitudes $|Z_1(t)|$ em 2003, sendo $h = 0,8$ para a série original (círculos) e $h = 0,5$ para a sua versão embaralhada (quadrados). **(c)** O análogo de (a) para a série dos sinais $\text{sign}[Z_1(t)]$ em 2003, sendo $h = 0,4$ para a série original (círculos) e $h = 0,5$ para a sua versão embaralhada (quadrados). Em (a), (b) e (c) as curvas $F(s)$ para as versões embaralhadas foram verticalmente deslocadas para melhor visualização. **(d)** Evolução temporal (anual) dos expoentes h para a série $D(t)$ (círculos), $|Z_1(t)|$ (quadrados) e $\text{sign}[Z_1(t)]$ (triângulos), calculados no período 1991-2003.

sugerem que grandes e positivos incrementos no índice Dst são mais provavelmente seguidos por grandes e negativos incrementos. De maneira similar, pequenos e negativos incrementos são mais provavelmente seguidos por pequenos e positivos incrementos e assim por diante.

7.4 Multifractalidade

Investigamos também a presença de multifractalidade na série $D(t)$ utilizando o método mf-DFA[30], descrito no capítulo 2. Especificamente, calculamos a função de flutuação generalizada $F_m(s)$ para escalas de tempo no intervalo $15 \leq n \leq 1080$ e para valores de m no intervalo $-5 \leq m \leq 5$. Em analogia com a eq. 7.7, verificamos que

$$F_m(s) \sim s^{h(m)+1}. \quad (7.8)$$

Na fig. 7.5a, vemos $F_m(s)$ para valores particulares de m , calculados para a série $D(t)$ do ano 2003. Na fig. 7.5b, vemos o espectro multifractal $\tau(m) = m h(m) - 1$, calculado para a série $D(t)$ em 2003. Observe que $\tau(m)$ é claramente uma função não linear de m , indicando comportamento multifractal. Por outro lado, $\tau(m)$ é praticamente linear para uma versão embaralhada da série original. Isso indica que o comportamento multifractal da série Dst origina-se basicamente de suas propriedades de correlação, e não da forma da distribuição.

Verificamos que o análogo das figs. 7.5a e 7.5b vale para cada um dos anos no período 1991-2003. Para visualizarmos o grau de multifractalidade no índice Dst em cada um destes anos, definimos a função $\Theta = H_</math>/ $H_>$, sendo $H_<$ a inclinação da função $\tau(m)$ para $m < 0$ e $H_>$ sendo a inclinação da função $\tau(m)$ para $m > 0$. Por definição, sinais monofractais são caracterizados por $\Theta = 1$. Na fig. 7.5c, vemos a função $\Theta(t)$, no período 1991-2003. Observe que $\Theta(t) > 1$ indicando que $\tau(m)$ é não linear em todos os anos do período considerado.$

Ademais, investigamos propriedades multifractais da série $D(t)$ em períodos de baixa e alta atividade geomagnética. Para isso, selecionamos o período de dois meses com menor valor médio de $D(t)$ (o mais calmo) e o período de dois meses com maior valor médio de $D(t)$ (o mais agitado) entre os anos 1991 e 2003. Calculamos a função de flutuação generalizada $F_m(s)$ para escalas de tempo no intervalo $15 \leq s \leq 96$ e para $-5 \leq m \leq 5$. Na fig. 7.5d, vemos os expoentes $\tau(m)$ para ambos os períodos. Observe

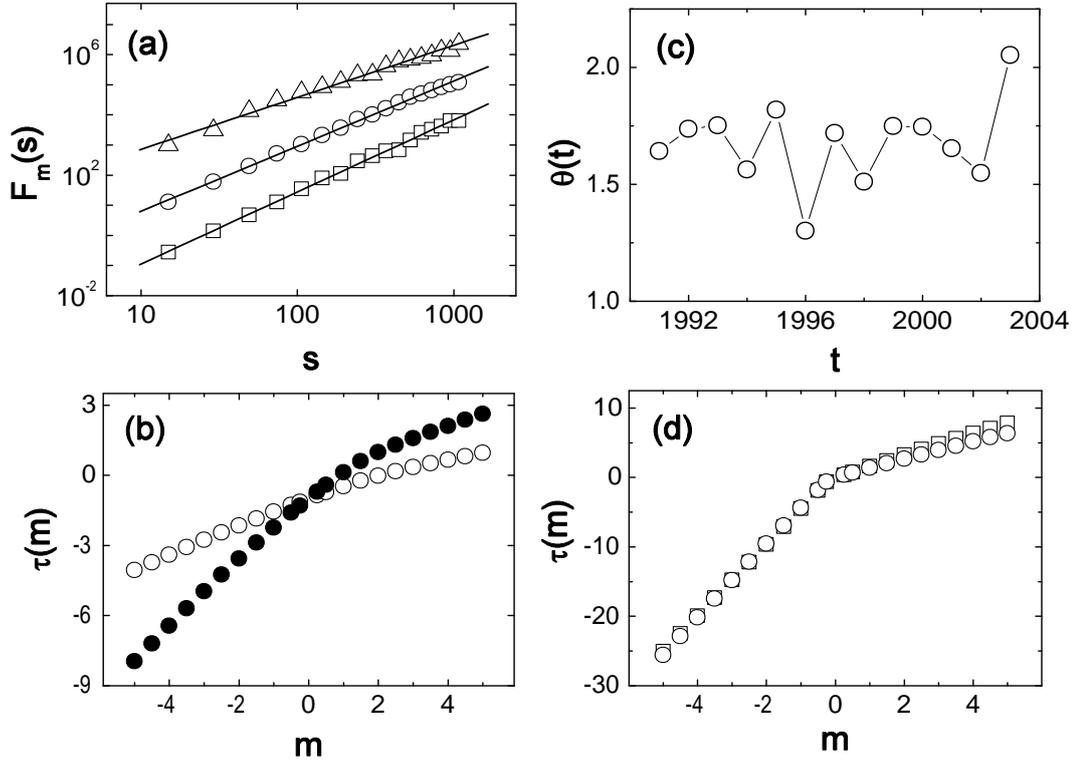


Figura 7.5: Multifractalidade no índice Dst. **(a)** Função de flutuação generalizada, $F_m(s)$, obtida via método mf-DfA de segunda ordem, para a série $D(t)$ em 2003. As linhas sólidas representam leis de potência, dadas pela eq. (7.8), com $h(-5) = 1,4$ (quadrados), $h(0,5) = 1,2$ (círculos) e $h(5) = 0,7$ (triângulos). Os expoentes referem-se ao intervalo $15 \leq s \leq 1080$. **(b)** Espectro multifractal $\tau(m) = m h(m) - 1$ para a série $D(t)$ em 2003 (círculos fechados) e para sua versão embaralhada (círculos abertos). Observe o comportamento fortemente não linear de $\tau(m)$ para a série original em contraste com o comportamento quase linear de $\tau(m)$ para sua versão embaralhada. **(c)** Função $\Theta(t)$ calculada no período 1991-2003. Note que $\Theta(t) > 1$ indicando que $\tau(m)$ é não linear em todos os anos do período considerado. Em particular, $\langle \Theta \rangle \simeq 1,7$. **(d)** Espectro multifractal $\tau(m)$ calculado para o período de dois meses mais calmo (círculos) e o período de dois meses mais agitado (quadrados) entre 1991 e 2003. Em ambos os casos o comportamento de $\tau(m)$ é fortemente não linear.

que $\tau(m)$ exibe praticamente o mesmo comportamento em ambos os casos, indicando que a série continua sendo multifractal mesmo em situações de baixa e alta atividade geomagnética.

Esses resultados são consistentes com evidências que indicam a presença de multifractalidade em outros sinais tipos de sinais geomagnéticos[135, 69].

7.5 Características universais

Dentre os resultados obtidos da análise da dinâmica do índice Dst, temos que: *i*) a distribuição dos incrementos consecutivos z_1 pode ser aproximada por uma distribuição de Lévy, com $\psi \simeq 1,7$, para pequenos valores de z_1 ; *ii*) a distribuição dos incrementos $p(z_n)$ exibe invariância de escala, preservando a mesma forma funcional para diferentes valores de n ; *iii*) as correlações temporais são descritas por expoentes $h \simeq 1,0$ para o índice Dst, $h \simeq 0,75$ para a série das magnitudes e $h \simeq 0,4$ para a série dos sinais; *iv*) o índice Dst exibe espectro multifractal $\tau(m)$ não linear, indicando a presença de multifractalidade; *v*) períodos de baixa e alta atividade geomagnética exibem propriedades multifractais similares; *vi*) as correlações de longo alcance e a multifractalidade são estáveis durante o período considerado.

Para comparação, considere os seguintes resultados análogos observados na dinâmica dos intervalos de tempo entre batimentos cardíacos: *i*) a distribuição dos incrementos consecutivos pode ser aproximada por uma distribuição de Lévy, com $\psi \simeq 1,7$ [91]; *ii*) a distribuição dos incrementos exibe invariância de escala, preservando a mesma forma funcional para diferentes escalas de tempo[95, 96]; *iii*) as correlações temporais são descritas por expoentes $h \simeq 1,0$ para os intervalos de tempo entre batimentos, $h \simeq 0,75$ para a série das magnitudes e $h \simeq 0,4$ para a série dos sinais[92]; *iv*) a série dos intervalos de tempo entre batimentos exibe um espectro multifractal $\tau(m)$ não linear, indicando a presença de multifractalidade[93]; *v*) situações de repouso e atividade física exibem propriedades multifractais similares[94]; *vi*) as correlações de longo alcance e a multifractalidade são estáveis para diferentes indivíduos.

Alguns resultados relacionados aos sinais em consideração são resumidos na tab. 7.1. Observe que a dinâmica do índice Dst exibe padrões surpreendentemente similares aos observados na dinâmica dos batimentos cardíacos.

Tabela 7.1 - Dinâmica do índice Dst em comparação com a dinâmica dos intervalos de tempo entre batimentos cardíacos de pessoas saudáveis.

	$p(z)$	ψ	$F(n)$	h_{orig}	h_{mag}	h_{sin}	$\tau(m)$
índice Dst	Lévy	1,7	$\sim n^{h+1}$	1,0	0,75	0,4	não linear
batimentos cardíacos	Lévy	1,7	$\sim n^{h+1}$	1,0	0,75	0,4	não linear

Capítulo 8

Conclusão

Ao aplicarmos conceitos e métodos de física estatística no estudo da circulação de revistas, das citações de periódicos científicos, de atividades religiosas e de um sinal geomagnético, verificamos que tais sistemas exibem características em comum, as quais podem ser agrupadas sobre os pilares da invariância de escala e da universalidade.

Observamos a presença de invariância de escala *i*) no comportamento assintótico do tipo lei de potência das distribuições de tamanho, com expoente $\alpha \simeq 2,5$ para a circulação, $\alpha \simeq 2,7$ para o fator de impacto e $\alpha \simeq 2,0$ para o número de membros e demais atividades religiosas; *ii*) no decaimento do tipo lei de potência do desvio padrão das taxas de crescimento com o tamanho, com expoente $\beta \simeq 0,22$ para o fator de impacto, $\beta \simeq 0,16$ para número de membros e total de horas, $\beta \simeq 0,13$ para a média de cursos dirigidos e $\beta \simeq 0,10$ para a média de pioneiros; *iii*) nas distribuições das taxas de crescimento do fator de impacto e de atividades religiosas, as quais exibem a mesma forma funcional para diferentes medidas de tamanho; *iv*) na distribuição dos incrementos no índice Dst, caracterizada pela mesma forma funcional para diferentes escalas de tempo, exibindo comportamento assintótico do tipo lei de potência, com expoente $\alpha \simeq 4$; *v*) nas correlações de longo alcance que governam a evolução temporal do índice Dst, caracterizadas pelos expoentes $h \simeq 1,0$ para a série original, $h \simeq 0,75$ para a magnitude dos incrementos consecutivos e $h \simeq 0,4$ para o sinal dos incrementos consecutivos; e *vi*) na multifractalidade do índice Dst, caracterizada por diversos expoentes generalizados $h(m)$ significativamente dependentes de m .

Verificamos também que os sistemas analisados apresentam padrões e características universais. Especificamente, certas propriedades estatísticas da circulação de revistas, do fator de impacto de periódicos científicos e de atividades religiosas exibem padrões

surpreendentemente similares aos observados na dinâmica do tamanho de firmas, do PIB de países e de atividades de pesquisa científica (veja as tabs. 4.1, 5.1, 5.2, 6.1, 6.2 e 6.3). Essas similaridades reforçam a hipótese de que diferentes tipos de organização exibem comportamento universal, sendo governadas por mecanismos similares. Deste modo, nossos resultados sugerem que os mecanismos que governam a dinâmica da circulação de revistas, do fator de impacto de periódicos científicos e de atividades religiosas das Testemunhas de Jeová, sejam análogos aos mecanismos que dirigem a dinâmica de atividades econômicas (tais como o tamanho de firmas) e atividades de pesquisa científica (tais como as pesquisas em universidades).

É bom lembrar também que a analogia entre atividades econômicas e atividades de pesquisa científica já era conhecida. Entretanto, pela primeira vez, mostramos que atividades religiosas pertencem a esta mesma classe de “universalidade” (veja as tabs. 6.1, 6.2 e 6.3). Em particular, no que se refere a atividades religiosas, há algo notável nas leis empíricas resumidas na tab. 6.1, que é o fato de que elas governam as taxas de crescimento de atividades religiosas de um conjunto diverso de países e ilhas - apesar de suas diferenças culturais, sociais e econômicas. Esses resultados também fazem lembrar o conceito de universalidade na dinâmica de sistemas complexos, em que o comportamento global não depende das particularidades do sistema.

A existência de mecanismos universais, independentes das particularidades do sistema em estudo, abre a possibilidade de se desenvolver modelos gerais que poderiam ser aplicados no estudo de sistemas distintos. Nesta direção, propomos dois modelos estocásticos: um no contexto da circulação de revistas (modelo I) e o outro no contexto do fator de impacto de periódicos científicos (modelo II). Ambos os modelos são generalizações do modelo de Gibrat para a dinâmica de firmas e baseiam-se na idéia de que as organizações podem ser consideradas como unidades sem estrutura interna complexa, seguindo processos multiplicativos. Apesar de sua simplicidade, esses modelos são capazes de reproduzir aspectos importantes do comportamento observado. Esperamos que eles possam ser úteis no estudo da dinâmica de outros sistemas complexos. Uma perspectiva interessante seria desenvolver modelos gerais que levam em conta a estrutura interna complexa das organizações.

Ademais, verificamos que a dinâmica do índice Dst exhibe notáveis similaridades com a dinâmica dos intervalos de tempo entre batimentos cardíacos de pessoas saudáveis. Ambos os sinais apresentam invariância de escala na distribuição dos incrementos em diferentes escalas de tempo, correlações de longo alcance e multifractalidade. Em particular, ressaltamos que algumas destas similaridades são quantitativas (veja tab.

7.1). Além disso, vimos que períodos de baixa e alta atividade geomagnética exibem propriedades multifractais similares. De forma similar, os intervalos de tempo entre batimentos cardíacos de indivíduos em repouso e de indivíduos em atividade física exibem propriedades multifractais similares. Ademais, a estabilidade que observamos nas correlações de longo alcance e na multifractalidade do índice Dst para diferentes anos é consistente com a estabilidade destas medidas para diferentes indivíduos.

Estas similaridades sugerem que os mecanismos que governam a interação do campo magnético da Terra com o vento solar são similares aos mecanismos que regulam a taxa dos batimentos cardíacos de pessoas saudáveis. Além disso, nossos resultados indicam a possibilidade de que o recente progresso no estudo da dinâmica dos batimentos cardíacos possa ser de ajuda no estudo dos mecanismos complexos que governam a atividade geomagnética. Há a necessidade de estudos adicionais para verificar até onde vai esta analogia entre sistemas tão diferentes. De qualquer forma, nossos resultados provêm um guia para desenvolver modelos mais realísticos para os processos que controlam a atividade geomagnética. Como perspectiva, desejamos desenvolver um modelo estocástico que seja capaz de reproduzir certos aspectos da dinâmica do índice Dst.

Apêndice A

Exemplos de distribuições

A função densidade de probabilidade associada a uma variável aleatória pode assumir diversas formas. A seguir, apresentaremos uma breve descrição de algumas distribuições, focalizando as que serão utilizadas ou mencionadas neste trabalho.

A.1 Distribuições Gaussiana e de Laplace

A distribuição normal, também chamada distribuição Gaussiana, é definida como

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right], \quad (\text{A.1})$$

onde μ é a média e σ^2 a variância de x . A eq. (A.1) apresenta-se na forma de parábolas na escala mono-log (fig. A.1a).

Por sua vez, a distribuição de Laplace (ou exponencial dupla)

$$p(x) = \frac{1}{\sqrt{2}\sigma} \exp\left[-\frac{\sqrt{2}|x - \mu|}{\sigma}\right], \quad (\text{A.2})$$

com parâmetros μ e σ , exibe um formato triangular na escala mono-log (fig. A.1b) e, por este motivo, tem sido chamada de distribuição em forma de tenda. Observe que, enquanto a distribuição Gaussiana é expressa em termos das diferenças quadráticas da média, a distribuição de Laplace é expressa em termos das diferenças absolutas da média. Conseqüentemente, a distribuição de Laplace decresce mais lentamente que a

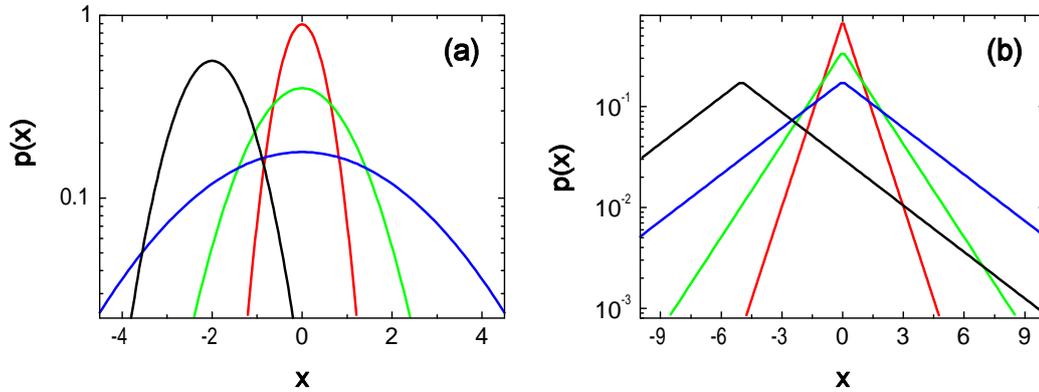


Figura A.1: **(a)** Distribuição Gaussiana (na escala mono-log) para valores particulares de μ e σ^2 . (cor, μ , σ^2): (vermelho, 0, 0.2), (verde, 0, 1.0), (azul, 0, 5.0), (preto, -2, 0.5). **(b)** Distribuição de Laplace (na escala mono-log) para valores particulares de μ e σ . (cor, μ , σ): (vermelho, 0, 1), (verde, 0, 2), (azul, 0, 4), (preto, -5, 4).

Gaussiana.

A.2 Leis de potência assintóticas

Como já comentado, distribuições do tipo lei de potência são características bem conhecidas de um grande número de sistemas complexos. Entretanto, a distribuição $p(x) \propto x^{-\alpha}$, com $\alpha > 0$, diverge para $x \rightarrow 0$ e não é normalizável no intervalo de zero a infinito. Em certos casos, estas leis de potência podem ser entendidas como o comportamento assintótico de distribuições mais gerais, tais como a distribuição de Lévy[31, 32] e distribuições que derivam da mecânica estatística não extensiva[33, 34, 8, 9]. No que segue, descreveremos brevemente algumas destas distribuições, destacando nossas recentes incursões neste tema.

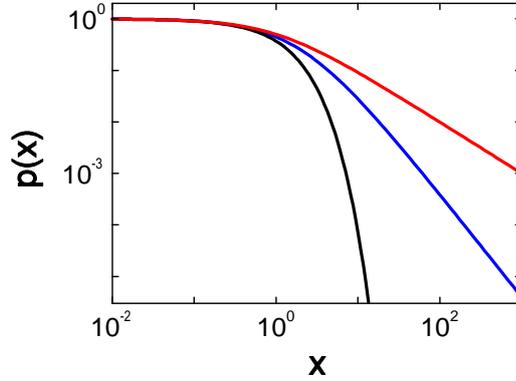


Figura A.2: Distribuição q -exponencial, dada pela eq. A.3 com $a = 1$, para $q = 1$ (preto), $q = 1,5$ (azul) e $q = 2,0$ (vermelho).

A.2.1 Distribuição q -exponencial

Considere, por exemplo, a distribuição q -exponencial definida como

$$p(x) = a \exp_q \left(-\frac{x}{b} \right) = a \left[1 + (q-1) \frac{x}{b} \right]^{-1/(q-1)}, \quad (\text{A.3})$$

sendo a e b constantes positivas. Se $q = 1$, a q -exponencial se reduz à exponencial usual, $p(x) = a \exp(-x/b)$, visto que $\exp_q(-x) \rightarrow \exp(-x)$ no limite $q \rightarrow 1$. No caso $q < 1$, define-se $p(x) = 0$ para $1 + (q-1)x/b < 0$. Observe que $\int_0^\infty p(x) dx = 1$, com $a = (2-q)/b$.

Se $q > 1$, a q -exponencial exibe um comportamento assintótico do tipo lei de potência,

$$p(x) \sim x^{-1/(q-1)}. \quad (\text{A.4})$$

Na fig. A.2, vemos a eq. (A.3) para valores típicos de q .

A distribuição acumulada de uma q -exponencial é também uma q -exponencial. De fato,

$$p_c(x) = a' \exp_{q'} \left(-\frac{x}{b'} \right) = a' \left[1 + (q'-1) \frac{x}{b'} \right]^{-1/(q'-1)}, \quad (\text{A.5})$$

com $q' = 1/(2-q)$, $b' = b/(2-q)$ e $a' = ab/(2-q)$.

As q -exponenciais têm sido utilizadas recentemente, por exemplo, no estudo das dis-

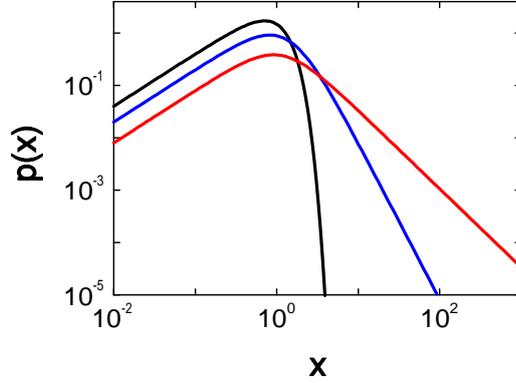


Figura A.3: Distribuição q -Weibull, dada pela eq. A.6 com $a = (2 - q)r^2/b^2$ e $b = 1$, para $q = 1$ e $r = 2$ (preto); $q = 1,5$ e $r = 2$ (azul); e $q = 1,8$ e $r = 2$ (vermelho).

tribuições de citações científicas[35], de gols no futebol[36], da população de cidades[37] e de votos em eleições[38].

A.2.2 Distribuição q -Weibull

Recentemente, propomos uma generalização da distribuição de Weibull, a q -Weibull, definida como[39]

$$p(x) = ax^{r-1} \exp_q \left(-\frac{x^r}{b} \right) = ax^{r-1} \left[1 + (q-1) \frac{x^r}{b} \right]^{-1/(q-1)}, \quad (\text{A.6})$$

com a , b e r constantes positivas (veja fig. A.3). No limite $q \rightarrow 1$, a eq. (A.6) reduz-se à distribuição de Weibull usual. No caso $q < 1$, definimos $p(x) = 0$ para $x \geq [b/(1-q)]^{1/r}$. Note que $\int_0^\infty p(x)dx = 1$, com $a = (2 - q)r^2/b^2$.

A distribuição q -Weibull, com $q > 1$, exhibe comportamento assintótico do tipo lei de potência para ambos os limites: pequenos e grandes valores de x . Especificamente,

$$p(x) \sim x^{-\xi}, \quad (\text{A.7})$$

com $\xi = -(r-1)$ para x pequeno e $\xi = r[(2-q)/(q-1)] + 1$ para x grande. Na fig. A.3, vemos a eq. (A.6) para valores típicos de q e r .

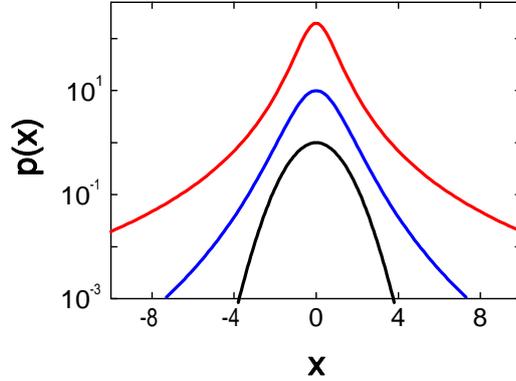


Figura A.4: Distribuição q -Gaussiana, dada pela eq. A.8 com $b = 5 - 3q$, para $q = 1$ (preto); $q = 1,3$ (azul); e $q = 1,5$ (vermelho). As distribuições foram deslocadas verticalmente para melhor visualização.

Recentemente, utilizamos a q -Weibull para descrever a distribuição do comprimento de rodovias[39] e a distribuição da tensão de quebra dielétrica em óxidos[40].

A.2.3 Distribuição q -Gaussiana

A distribuição q -Gaussiana é definida como

$$p(x) = a \exp_q \left(-\frac{x^2}{b} \right) = a \left[1 + (q-1) \frac{x^2}{b} \right]^{-1/(q-1)}, \quad (\text{A.8})$$

com a e b constantes positivas. No limite $q \rightarrow 1$, a eq. A.8 reduz-se à distribuição Gaussiana usual. Se $b = 5 - 3q$, com $q < 5/3$, a distribuição q -Gaussiana tem variância finita (unitária).

Se $q > 1$, as caudas da distribuição q -Gaussiana decrescem como leis de potência

$$p(|x|) \sim |x|^{-2/(q-1)}. \quad (\text{A.9})$$

Na fig. A.4, vemos a eq. (A.8) para valores típicos de q .

A distribuição q -Gaussiana tem sido aplicada, por exemplo, no estudo de séries econômicas tais como taxas de câmbio[41] e índice de ações[8, 42]. Em particular,

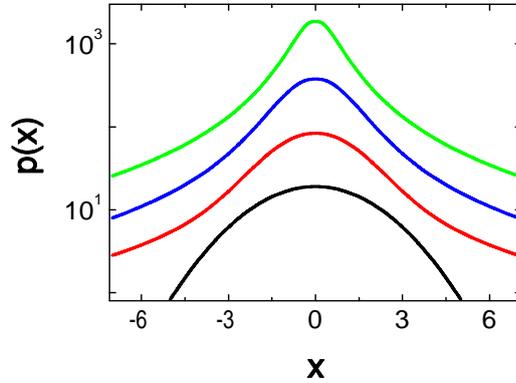


Figura A.5: Distribuição simétrica de Lévy, dada pela eq. A.10 (com $\rho = 0,8$), para $\psi = 2$ (preto), $\psi = 1,7$ (vermelho), $\psi = 1,4$ (azul) e $\psi = 1$ (verde).

recentemente aplicamos a função q -Gaussiana num modelo de redes de polímeros, fornecendo uma melhor aproximação para redes de tamanho finito[43].

A.2.4 Distribuição simétrica de Lévy

A distribuição simétrica de Lévy, com índice ψ e fator de escala ρ , é definida como

$$p(x) = \frac{1}{\pi} \int_0^{\infty} \exp(-\rho k^{\psi}) \cos(kx) dk. \quad (\text{A.10})$$

Trata-se de uma distribuição normalizada, cuja forma analítica é conhecida, em termos de funções elementares, apenas para $\psi = 2$ (Gaussiana) e $\psi = 1$ (Lorentziana).

O comportamento assintótico da distribuição simétrica de Lévy, para grandes valores de $|x|$, é

$$p(|x|) \sim |x|^{-(1+\psi)}. \quad (\text{A.11})$$

Na fig. A.5, vemos exemplos da eq. (A.10) para valores particulares de ψ .

Bibliografia

- [1] F. Mallamace and H. E. Stanley, Proceedings of the International School of Physics “Enrico Fermi”, Course CLV, *The Physics of Complex Systems*, (IOS Press, Amsterdam, 2004).
- [2] R. N. Mantegna and H. E. Stanley, *An Introduction to Econophysics - Correlations and Complexity in Finance*, (Cambridge University Press, Cambridge, 2000).
- [3] J. Voit, *The Statistical Mechanics of Financial Markets*, (Springer-Verlag, Berlin, 2001).
- [4] W. Paul and J. Baschnagel, *Stochastic processes - From Physics to Finance* (Springer-Verlag, Berlin, 1999).
- [5] J. B. Bassingthwaite, L. S. Liebovitch and B. J. West, *Fractal Physiology* (Oxford University Press, New York, 1994).
- [6] M. Small, *Applied Nonlinear Time Series Analysis - Applications in Physics, Physiology and Finance* (World Scientific Publishing, Singapore, 2005).
- [7] N. Stergiou, *Innovative Analyses of Human Movement - Analytical Tools for Human Movement Research* (Human Kinetics, Champaign, 2004).
- [8] M. Gell-Mann and C. Tsallis, *Nonextensive entropy - Interdisciplinary Applications*, (Oxford University Press, New York, 2004).
- [9] S. Abe and Y. Okamoto, *Nonextensive Statistical Mechanics and Its Applications* (Springer-Verlag, New York, 2001).
- [10] C.-K. Peng S. V. buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons and H. E. Stanley, Long-range correlations in nucleotide sequences, *Nature* **356**, 168 (1992).

- [11] E. K-Bunde, A. Bunde, S. Havlin, H. E. Roman, Y. Goldreich and H. J. Schellnhuber, Indication of a universal persistence law governing atmospheric variability, *Physical Review Letters* **81**, 729 (1998).
- [12] P. Ch. Ivanov, M. G. Rosenblum, C.-K. Peng, J. Mietus, S. Havlin, H. E. Stanley and A. L. Goldberger, Scaling behavior of heartbeat intervals obtained by wavelet-based time series analysis, *Nature* **383**, 323 (1996).
- [13] R. N. Mantegna and H. E. Stanley, Scaling behavior in the dynamics of an economic index, *Nature* **376**, 46 (1995).
- [14] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner and Y. Dodge, Turbulent cascades in foreign exchange markets, *Nature* **381**, 767 (1996).
- [15] T. H. Keitt and H. E. Stanley, Dynamics of North American breeding bird populations, *Nature* **393**, 257 (1998).
- [16] C. J. Rhodes and R. M. Anderson, Power laws governing epidemics in isolated populations, *Nature* **381**, 600 (1996).
- [17] D. Brockmann, L. Hufnagel and T. Geisel, The scaling laws of human travel, *Nature* **439**, 462 (2006).
- [18] H. E. Stanley, L. A. N. Amaral, S. V. Buldyrev, A. L. Goldberger, S. Havlin, H. Leschhorn, P. Maass, H. A. Makse, C.-K. Peng, M. A. Salinger, M. H. R. Stanley and G. M. Viswanathan, Scaling and universality in animate and inanimate systems, *Physica A* **231**, 20 (1996).
- [19] H. E. Stanley, L. A. N. Amaral, P. Gopikrishnan, P.-Ch. Ivanov, T. H. Keitt and V. Plerou, Scale invariance and universality: organizing principles in complex systems, *Physica A* **281**, 60 (2000).
- [20] S. Picoli Jr., R. S. Mendes and L. C. Malacarne, Statistical properties of the circulation of magazines and newspapers, *Europhysics Letters* **72**, 865 (2005).
- [21] S. Picoli Jr., R. S. Mendes, L. C. Malacarne and E. K. Lenzi, Scaling behavior in the dynamics of citations to scientific journals, *Europhysics Letters* **75**, 673 (2006).
- [22] S. Picoli Jr. and R. S. Mendes, Universal features in the growth dynamics of religious activities, submitted (2007).

- [23] S. Picoli Jr., R. S. Mendes, L. C. Malacarne and A. R. R. Papa, Similarities between the dynamics of geomagnetic signal and of heartbeat intervals, accepted in *Europhysics Letters* (2007).
- [24] W. Feller, *An introduction to Probability Theory and its Applications* (Wiley, London, 1950).
- [25] N. G. Van Kampen, *Stochastic processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).
- [26] J. Feder, *Fractals* (Plenum Press, New York, 1988).
- [27] T. Vicsek, *Fractal Growth Phenomena* (World Scientific, Singapore, 1993).
- [28] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley and A. L. Goldberger, Mosaic organization of DNA nucleotides, *Physical Review E* **49**, 1685 (1994).
- [29] S. V. Buldyrev, A. L. Goldberger, S. Havlin, R. N. Mantegna, M. E. Matsu, C.-K. Peng, M. Simons and H. E. Stanley, Long-range correlation-properties of coding and noncoding DNA-sequences - Genbank analysis, *Physical Review E* **51**, 5084 (1995).
- [30] J. W. Kantelhardt, S. A. Zschiegner, E. K.-Bunde, S. Havlin, A. Bunde and H. E. Stanley, Multifractal detrended fluctuations analysis of nonstationary time series, *Physica A* **316**, 87 (2002).
- [31] P. Lévy, *Calcul des probabilités* (Gauthier-Villars, Paris, 1925).
- [32] A. Ya. Khintchine and P. Lévy, Sur les loi stables, *C. R. Acad. Sci. Paris* **202**, 374 (1936).
- [33] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, *Journal of Statistical Physics* **52**, 479 (1988).
- [34] C. Tsallis, R. S. Mendes and A. R. Plastino, The role of constraints within generalized nonextensive statistics, *Physica A* **261**, 534 (1998).
- [35] C. Tsallis and M. P. de Albuquerque, Are citations of scientific papers a case of nonextensivity? *European Physical Journal B* **13**, 777 (2000).

- [36] L. C. Malacarne and R. S. Mendes, Regularities in football goal distributions, *Physica A* **286**, 391 (2000).
- [37] L. C. Malacarne, R. S. Mendes and E. K. Lenzi, q -exponential distribution in urban agglomeration, *Physical Review E* **65**, 017106 (2002).
- [38] M. L. Lyra, U. M. S. Costa, R. N. Costa Filho and J. S. Andrade Jr., Generalized Zipf's law in proportional voting processes, *Europhysics Letters* **62**, 131 (2003).
- [39] S. Picoli Jr., R. S. Mendes and L. C. Malacarne, q -exponential, Weibull and q -Weibull distributions: an empirical analysis, *Physica A* **324**, 678 (2003).
- [40] U. M. S. Costa, V. N. Freire, L. C. Malacarne, R. S. Mendes, S. Picoli Jr., E. A. de Vasconcelos and E. F. da Silva Jr., An improved description of the dielectric breakdown in oxides based on a generalized Weibull distribution, *Physica A* **361**, 209 (2006).
- [41] F. M. Ramos, C. Rodrigues Neto, R. R. Rosa, L. D. Abreu Sá and M. J. A. Bolzam, Generalized thermostistical description of intermittency and nonextensivity in turbulence and financial markets, *Nonlinear Analysis* **23**, 3521 (2001).
- [42] F. Michael and M. D. Johnson, Financial market dynamics, *Physica A* **320**, 525 (2003).
- [43] L. C. Malacarne, R. S. Mendes, E. K. Lenzi, S. Picoli Jr. and J. P. Dal Molin, A non-Gaussian model in polimeric network, *The European Physical Journal E* **20**, 395 (2006).
- [44] H. E. Hurst, R. P. Black and Y. M. Simaika, *Long-term storage. An experimental study* (Constable, London, 1965).
- [45] J. Beran, *Statistics for Long-Memory Processes* (Chapman & Hall/CRC, New York, 1994).
- [46] J. F. Muzy, E. Bacry and A. Arneodo, Wavelets and multifractal formalism for singular signals - application to turbulence data, *Physical Review Letters* **67**, 3515 (1991).
- [47] R. N. Mantegna, S. V. Buldyrev, A. L. Goldberger, S. Havlin, C.-k. Peng, M. Simons and H. E. Stanley, Linguistic features of nonconding DNA-sequences, *Physical Review Letters* **73**, 3169 (1994).

- [48] P. Talkner and R. O. Weber, Power spectrum and detrended fluctuation analysis: Application to daily temperatures, *Physical Review E* **62**, 150 (2000).
- [49] M. Ausloos and K. Ivanova, Dynamical model and nonextensive statistical mechanics of a market index on large time windows, *Physical Review E* **68**, 046122 (2003).
- [50] J. M. Hausdorff, S. L. Mitchell, R. Firtion, C.-K. Peng, M. E. Cudkowicz, J. Y. Wei and A. L. Goldberger, Altered fractal dynamics of gait: Reduced stride-interval correlations with aging and Huntington's disease , *Journal of Applied Physiology* **82**, 262 (1997).
- [51] A. Bunde, S. Havlin, J. W. Kantelhardt, T. Penzel, J. H. Peter and K. Voigt, Correlated and uncorrelated regions in heart-rate fluctuations during sleep, *Physical Review Letters* **85**, 3736 (2000).
- [52] K. Ivanova, M. Ausloos, E. E. Clothiaux and T. P. Ackerman, Break-up of stratus cloud structure predicted from non-Brownian motion liquid water and brightness temperature fluctuations, *Europhysics Letters* **52**, 40 (2000).
- [53] S. Bahar, J. W. Kantelhardt, A. Neiman, H. H. A. Rego, D. F. Russell, L. Wilkens, A. Bunde and F. Moss, Long-range temporal anti-correlations in paddlefish electroreceptors , *Europhysics Letters* **56**, 454 (2001).
- [54] L. A. N. Amaral, D. J. B. Soares, L. R. da Silva, L. S. Lucena, M. Saito, H. Kumano, N. Aoyagi and Y. Yamamoto, Power law temporal auto-correlations in day-long records of human physical activity and their alteration with disease, *Europhysics Letters* **66**, 448 (2004).
- [55] L. Telesca and V. Lapenna, Magnitude and sign scaling in power-law correlated geoelectrical time series measured in southern Italy, *Natural Hazards and Earth System Sciences* **4**, 669 (2004).
- [56] R. B. Govindan and H. Kantz, Long-term correlations and multifractality in surface wind speed, *Europhysics Letters* **68**, 184 (2004).
- [57] J. A. Wanliss, Nonlinear variability of SYM-H over two solar cycles, *Earth Planets Space* **56**, e13 (2004).

- [58] L. M. Stadler, B. Sepiol B, B. Pfau, J. W. Kantelhardt, R. Weinkamer and G. Vogl, Detrended fluctuation analysis in x-ray photon correlation spectroscopy for determining coarsening dynamics in alloys, *Physical Review E* **74**, 041107 (2006).
- [59] K. Hu, P.-Ch. Ivanov, Z. Chen, P. Carpena and H. E. Stanley, Effect of trends on detrended fluctuation analysis, *Physical Review E* **64**, 011114 (2001).
- [60] J. Kantelhardt, Detecting long-range correlations with detrended fluctuation analysis, *Physica A* **295**, 441 (2001).
- [61] Z. Chen, P.-Ch. Ivanov, K. Fu and H. E. Stanley, Effect of nonstationarities on detrended fluctuation analysis, *Physical Review E* **65**, 041107 (2002).
- [62] J. Kantelhardt, Multifractality of river runoff and precipitation: comparison of fluctuation analysis and wavelet methods, *Physica A* **330**, 240 (2003).
- [63] P. Oswiecimka, J. Kwapien and S. Drozd, Wavelet versus detrended fluctuation analysis of multifractal structures, *Physical Review E* **74**, 016103 (2006).
- [64] K. Matia, Y. Ashkenazy and H. E. Stanley, Multifractal properties of price fluctuations of stocks and commodities, *Europhysics Letters* **61**, 422 (2003).
- [65] P. Oswiecimka, J. Kwapien and S. Drozd, Multifractality in the stock market: price increments versus waiting times, *Physica A* **347**, 626 (2005).
- [66] J. Kwapien, P. Oswiecimka and S. Drozd, Components of multifractality in high-frequency stock returns, *Physica A* **350**, 466 (2005).
- [67] L. Telesca, G. Colangelo, V. Lapenna and M. Macchiato, Fluctuation dynamics in geoelectrical data: an investigation by using multifractal detrended fluctuation analysis, *Physics Letters A* **332**, 398 (2004).
- [68] Y. Ida, M. Hayakawa, A. Adalev and K. Gotoh, Multifractal analysis for the ULF geomagnetic data during the 1993 Guam earthquake, *Nonlinear Processes in Geophysics* **12**, 157 (2005).
- [69] G. Currenti, C. del Negro, V. Lapenna and L. Telesca, Multifractality in local geomagnetic field at Etna volcano, Sicily (southern Italy), *Natural Hazards and Earth System Sciences* **5**, 555 (2005).

- [70] L. Telesca, V. Lapenna and M. Macchiato, Multifractal fluctuations in seismic interspike series, *Physica A* **354**, 629 (2005).
- [71] M. Baumert, V. Baier and A. Voss, Long-term correlations and fractal dimension of beat-to-beat blood pressure dynamics, *Fluctuation and Noise Letters* **5**, L549 (2005).
- [72] C. Varotsos, Power-law correlations in column ozone over Antarctica, *International Journal of Remote Sensing* **26**, 3333 (2005).
- [73] R. Barille and P. LaPenna, Multifractality of laser beam spatial intensity in a turbulent medium, *Applied Optics* **45**, 3331 (2006).
- [74] M. S. Movahed, G. R. Jafari, F. Ghasemi, S. Rahvar and M. R. R. Tabar, Multifractal detrended fluctuations analysis of sunspot time series, *Journal of Statistical Mechanics - Theory and Experiment* p02003 (2006).
- [75] R. Gibrat, *Les Inégalités Economiques* (Sirey, Paris, 1931).
- [76] Y. Ijiri and H. A. Simon, *Skew Distributions and the Sizes of Business Firms* (North-Holland, New York, 1977).
- [77] B. H. Hall, The relationship between firm size and firm growth in the US manufacturing sector, *J. Indust. Econ.* **35**, 583 (1987).
- [78] J. Sutton, Gibrat's legacy, *J. Econ. Lit.* **35**, 40 (1997).
- [79] M. H. R. Stanley, L. A. N. Amaral, S. V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger and H. E. Stanley, Scaling behavior in the growth of companies, *Nature* **379**, 804 (1996).
- [80] L. A. N. Amaral, S. V. Buldyrev, S. Havlin, P. Maass, M. A. Salinger, H. E. Stanley and M. H. R. Stanley, Scaling behavior in economics: The problem of quantifying company growth, *Physica A* **244**, 1 (1997).
- [81] Y. Fujiwara, C. Di Guilmi, H. Aoyama, M. Gallegati and W. Souma, Do Pareto-Zipf and Gibrat laws hold true? An analysis with European firms, *Physica A* **335**, 197 (2004).

- [82] Y. Lee, L. A. N. Amaral, D. Canning, M. Meyer and H. E. Stanley, Universal features in the growth dynamics of complex organizations, *Physical Review Letters* **81**, 3275 (1998).
- [83] V. Plerou, L. A. N. Amaral, P. Gopikrishnan, M. Meyer and H. E. Stanley, Similarities between the growth dynamics of university research and of competitive economic activities, *Nature* **400**, 433 (1999).
- [84] K. Matia, L. A. N. Amaral, M. Luwel, H. F. Moed and H. E. Stanley, Scaling phenomena in the growth dynamics of scientific output, *Journal of the American Society for Information Science and Technology* **56**, 893 (2005).
- [85] F. Havemann, M. Heinz and R. W.-Dobler, Firm-like behavior of journals? Scaling properties of their output and impact growth dynamics, *Journal of the American Society for Information Science and Technology* **56**, 3 (2005).
- [86] L. Liang, F. Havemann, M. Heinz and R. W.-Dobler, Structural similarities between science growth dynamics in China and in western countries, *Scientometrics* **66**, 311 (2006).
- [87] L. A. N. Amaral, S. V. Buldyrev, S. Havlin, M. A. Salinger and H. E. Stanley, Power law scaling for a system of interacting units with complex internal structure, *Physical Review Letters* **80**, 1385 (1998).
- [88] D. F. Fu, F. Pammolli, S. V. Buldyrev, M. Riccaboni, K. Matia, K. Yamasaki and H. E. Stanley, The growth of business firms: Theoretical framework and empirical evidence, *Proceedings of the National Academy of Sciences of the United States of America* **102**, 18801 (2005).
- [89] D. Fu, S. V. Buldyrev, M. A. Salinger and H. E. Stanley, Percolation model for growth rates of aggregates and its application for business firm growth, *Phys. Rev. E* **74**, 036118 (2006).
- [90] M. Kobayashi and T. Musha, 1/f Fluctuation of heartbeat period, *IEEE Trans. Biomed. Eng.* **29**, 456 (1982).
- [91] C.-K. Peng, J. Mietus, J. M. Hausdorff, S. Havlin, H. E. Stanley and A. L. Goldberger, Long-range anticorrelations and non-Gaussian behavior of the heartbeat, *Physical Review Letters* **70**, 1343 (1993).

- [92] Y. Ashkenazy P.-Ch. Ivanov, S. Havlin, C.-K. Peng, A. L. Goldberger and H. E. Stanley, Magnitude and signal correlations in heartbeat fluctuations, *Physical Review Letters* **86**, 1900 (2001).
- [93] P.-Ch. Ivanov, L. A. N. Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik and H. E. Stanley, Multifractality in human heartbeat dynamics, *Nature* **399**, 461 (1999).
- [94] L. A. N. Amaral, P.-Ch. Ivanov, N. Aoyagi, I. Hidaka, S. Tomono, A. L. Goldberger, H. E. Stanley and Y. Yamamoto, Behavioral-independent features of complex heartbeat dynamics, *Physical Review Letters* **86**, 6026 (2001).
- [95] K. Kiyono, Z. R. Struzik, N. Aoyagi, S. Sakata, J. Hayano and Y. Yamamoto, Critical scale invariance in a healthy human heart rate, *Physical Review Letters* **93**, 178103 (2004).
- [96] K. Kiyono, Z. R. Struzik, N. Aoyagi, F. Togo and Y. Yamamoto, Phase transition in a healthy human heart rate, *Physical Review Letters* **95**, 058101 (2005).
- [97] P.-Ch. Ivanov, A. Bunde, L. A. N. Amaral, S. Havlin, J. Fritsch-Yelle, R. M. Baevsky, H. E. Stanley and A. L. Goldberger, Sleep-wake differences in scaling behavior of the human heartbeat: Analysis of terrestrial and long-term space flight data, *Europhysics Letters* **48**, 594 (1999).
- [98] J. W. Kantelhardt, Y. Ashkenazy, P.-Ch. Ivanov, A. Bunde, S. Havlin, T. Penzel, J. H. Peter and H. E. Stanley, Characterization of sleep stages by correlations in the magnitude and sign of heartbeat increments, *Phys. Rev. E* **65**, 051908 (2002).
- [99] R. Karasik, N. Sapir, Y. Ashkenazy, P.-Ch. Ivanov, I. Dvir, P. Lavie and S. Havlin, Correlation differences in heartbeat fluctuations during rest and exercise, *Phys. Rev. E* **66**, 062902 (2002).
- [100] P.-Ch. Ivanov, Z. Chen, K. Hu and H. E. Stanley, Multiscale aspects of cardiac control, *Physica A* **344**, 685 (2004).
- [101] T. J. P. Penna, P. M. C. de Oliveira, J. C. Sartorelli, W. M. Gonçalves and R. D. Pinto, Long-range anticorrelations and non-Gaussian behavior of a leaky faucet, *Phys. Rev. E* **52**, R2168 (1995).

- [102] K. Fukuda, L. A. N. Amaral and H. E. Stanley, Similarities between communication dynamics in the Internet and the autonomic nervous system, *Europhysics Letters* **62**, 169 (2003).
- [103] R. Krishnan and L. Soley, Controlling magazine circulation, *Journal of Advertising Research* **27**, 17 (1987).
- [104] J. K. Kalita and R. H. Ducoffe, A simultaneous-equation analysis of pricing, circulation, and advertising revenue for leading consumer magazines, *Journal of Media Economics* **8**, 1 (1995).
- [105] M. A. Koschat and W. P. Putsis, Audience characteristics and bundling: A hedonic analysis of magazine advertising rates, *Journal of Marketing Research* **39**, 262 (2002).
- [106] C. A. Depken, Audience characteristics and the price of advertising in a circulation industry: evidence from US magazines, *Information Economics and Policy* **16**, 179 (2004).
- [107] Audit Bureau of Circulations, www.accessabc.com, www.abc.org.uk.
- [108] BPA Worldwide, www.bpaww.com.
- [109] E. Garfield, Citation indexes for science - new dimension in documentation through association of ideas, *Science* **122**, 108 (1955).
- [110] E. Garfield, Citation analysis as a tool in journal evaluation - journals can be ranked by frequency and impact of citations for science policy studies, *Science* **178**, 471 (1972).
- [111] Scientific Thomson, scientific.thomson.com.
- [112] P. O. Seglen, The skewness of science, *Journal of the American Society for Information Science* **43**, 628 (1992).
- [113] A. J. Lotka, *Journal of the Washington Academy of Sciences* **16**, 317 (1926).
- [114] E. Garfield, I. H. Sher, New factors in evaluation of scientific literature through citation indexing, *American Documentation* **14**, 195 (1963).
- [115] D. J. de S. Price, Networks of scientific papers, *Science* **149**, 510 (1965).

- [116] The library association. Report of the *Association's Committee on Scientific Library Sciences* (London,1968).
- [117] R. J. Cole, S. Cole, Ortega hypothesis, *Science* **178**, 368 (1972).
- [118] S. Naranan, Bradfords law of bibliography of science, *Nature* **227**, 631 (1970).
- [119] W. Goffman, T. G. Morris, Bradfords law and library acquisitions, *Nature* **226**, 922 (1970).
- [120] Watchtower - Official Web Site of Jehovah's Witnesses, www.watchtower.org (2006).
- [121] C. R. Wah, An introduction to research and analysis of Jehovah's Witnesses: A view from the Watchtower, *Review of Religious Research* **43**, 161 (2001).
- [122] C. R. Wah, Jehovah's-Witnesses and the responsibility of religious freedom: The European experience, *Journal of Church and State* **43**, 579 (2001).
- [123] C. R. Wah, Jehovah's Witnesses and the Empire of the Sun: A clash of faith and religion during World-War-II, *Journal of Church and State* **44**, 45 (2002).
- [124] J. L. Dixon and M. G. Smalley, Jehovah's Witnesses: The surgical-ethical challenge, *Journal of the American Medical Association* **246**, 2471 (1981).
- [125] R. Stark and L. Iannaccone, Why the Jehovah's Witnesses Grow so Rapidly: A Theoretical Application, *Journal of Contemporary Religion* **12**, 133 (1997).
- [126] L. J. Lanzerotti, *Impacts of ionospheric / magnetospheric process on terrestrial science and technology* (North Holland Publishing Co., New York, 1979).
- [127] H. B. Garrett, *Space Systems and Their Interactions with Earth's Space Environment* (American Institute of Aeronautics and Astronautics, New York, 1980).
- [128] S. R. C. Malin and B. J. Srivatava, Correlation between heart attacks and magnetic activity, *Nature* **277**, 646 (1979).
- [129] S. Dimitrova, Relationship between human physiological parameters and geomagnetic variations of solar origin, *Advances in Space Research* **37**, 1251 (2006).
- [130] International Association of Geomagnetism and Aeronomy (IAGA) **40** (1991).

- [131] W. D. Gonzalez, J. A. Joselyn, Y. Kamide, H. W. Kroehl, G. Rostoker, B. T. Tsurutani and V. M. Vasyliunas, What is a geomagnetic storm, *Journal of Geophysical Research-Space Physics* **99**, 5771 (1994).
- [132] W. H. Campbell, Geomagnetic storms, Dst ring-current myth and lognormal distributions, *Journal of Atmospheric and Terrestrial Physics* **58**, 1171 (1996).
- [133] World Data Centre for Geomagnetism, <http://swdcdb.kugi.kyoto-u.ac.jp>.
- [134] H. L. Wei, S. A. Billings and M. Balikhin, Analysis of the geomagnetic activity of the Dst index and self-affine fractals using wavelet transforms, *Nonlinear Processes in Geophysics* **11**, 303 (2004).
- [135] Z. Voros, On multifractality of high-latitude geomagnetic fluctuations, *Annales Geophysicae - Atmospheres Hydrospheres and Space Sciences* **18**, 1273 (2000).