

PROCESSOS ESTOCÁSTICOS

Celia Anteneodo
Centro Brasileiro de Pesquisas Físicas

RESUMO

Praticamente qualquer sistema (físico ou não) está sujeito a complicadas influências que não podem ser inteiramente conhecidas. Estas influências, tipicamente associadas a um grande número de graus de liberdade envolvidos, impedem prever com precisão arbitrária o estado do sistema em cada instante. Assim, o resultado de uma experiência programada para medir um dado observável tem uma componente que flutua ao se repetir a experiência, mesmo sendo esta preparada em condições praticamente idênticas. Apesar dessa imprevisibilidade, que caracteriza o fenômeno como sendo aleatório, quando a experiência é repetida um grande número de vezes aparecem regularidades. Isto permite a formulação de leis matemáticas onde os conceitos de probabilidade e de dinâmica estocástica, objetos de estudo do presente curso, são as noções fundamentais.

Nos sistemas quânticos existe uma imprevisibilidade adicional além da que surge devido exclusivamente à natureza complicada que o sistema possa ter.

Neste curso nos restringiremos a analisar principalmente exemplos da mecânica clássica. Também serão consideradas aplicações transdisciplinares.

O objetivo do curso é oferecer uma visão mais conceitual do que formal dos fenômenos estocásticos, complementando a abordagem analítica com experiências computacionais simples. Nestas notas serão apresentados os elementos básicos. Algumas propriedades em vez de serem enunciadas serão colocadas como questões para o leitor. As respostas a estas questões serão em geral úteis em discussões seguintes.

1 Probabilidade

Um experimento *aleatório* (ou não-determinista) é tal que não é possível afirmar *a priori* o resultado que ocorrerá, podendo o resultado ser diferente mesmo ao se repetir o ensaio em condições praticamente inalteradas. Os resultados podem parecer erráticos nas primeiras tentativas, entretanto, após um grande número de repetições, aparecem regularidades.

Quando lidamos com fenômenos aleatórios, podemos porém conhecer, em geral, o conjunto dos possíveis resultados a serem observados ao realizar uma dada experiência. A partir de um modelo, podemos também atribuir aos resultados ou conjuntos de resultados possíveis, números que representem as suas chances de ocorrência. Estes números, não-negativos e somando um para todos os possíveis resultados excludentes, são denominados probabilidades. O modelo pode ser construído seja a partir da frequência de ocorrência observada em um grande número de experimentos passados (lei dos grandes números) ou teoricamente, *a priori*.

Para uma revisão dos conceitos fundamentais em probabilidade, remeto o leitor para o capítulo deste livro da Profa. Maria Eulália Vares. Contudo, nesta seção encontra-se um compêndio da terminologia básica (veja por exemplo ¹⁾).

1.1 Variáveis aleatórias

Dado um experimento aleatório \mathcal{E} , o *espaço amostral* S é o conjunto dos resultados possíveis (numéricos ou não). A cada evento A (qualquer subconjunto de S) pode ser associado um número real não-negativo $P(A)$ denominado probabilidade tal que $P(A \cup B) = P(A) + P(B)$, para eventos A e B mutuamente excludentes, e $P(S) = 1$.

Uma *variável aleatória* (VA) unidimensional X é uma função $S \mapsto S_X \subset \mathbb{R}$ que associa a cada elemento $s \in S$ um (único) número real $X(s)$. A variável X pode ser discreta (S_X finito ou infinito numerável) ou contínua (S_X infinito não-numerável). Em alguns casos nos referiremos a propriedades das variáveis

discretas e em outros das variáveis contínuas unidimensionais, entretanto a adaptação a outros casos é em geral imediata.

1.2 Distribuições de probabilidade

A distribuição de probabilidade de uma variável aleatória X cujo contradomínio (ou espaço amostral com relação a X) é $S_X = \{x_1, x_2, \dots\}$ (caso discreto) é dada pelo conjunto de pares (x_i, p_i) , $i = 1, 2, \dots$, onde os $p_i \equiv P(X = x_i)$, probabilidade de x_i , devem satisfazer I) $p_i \geq 0$ (não-negatividade) e II) $\sum_{i \geq 1} p_i = 1$ (normalização).

No caso contínuo, a distribuição de probabilidade de uma variável aleatória X é dada pela função f_X , chamada função *densidade* de probabilidade (FDP), tal que $f_X(x)dx$ representa a probabilidade $P(x \leq X \leq x + dx)$. Assim, a FDP permite calcular a probabilidade de que X se encontre dentro de um intervalo real $[a, b]$:

$$P(a \leq X \leq b) = \int_a^b dx f_X(x). \quad (1)$$

Analogamente ao caso discreto deve valer $f_X(x) \geq 0, \forall x$, e $\int_{-\infty}^{\infty} dx f_X(x) = 1$.

A FDP *acumulada* é

$$F_X(x) = \int_{-\infty}^x ds f_X(s) = P(X \leq x). \quad (2)$$

Inversamente, a densidade não-acumulada é obtida por diferenciação:

$$f_X(x) = \frac{\partial F_X}{\partial x}. \quad (3)$$

Q. 1: Quanto vale $P(X = a)$ no caso contínuo? Quanto vale $F_X(\infty)$?

Para calcular a FDP de $Y = g(X)$, uma função de uma variável aleatória cuja FDP é conhecida, podemos proceder como segue

$$f_Y(y) = \int_{-\infty}^{\infty} dx' \delta(g(x') - y) f_X(x'). \quad (4)$$

Para este procedimento é útil a seguinte propriedade da função delta de Dirac: $\delta(g(x') - y) = \sum_x \delta(x' - x)/|g'(x)|$, onde x são os zeros de $g(x') - y$.

Q. 2: Seja X uniformemente distribuída no intervalo $[0, 1]$, ache $Y(X)$ tal que sua densidade seja exponencial $f_Y(y) = \alpha e^{-\alpha y}$, $y \geq 0$.

A FDP *conjunta* de duas (ou mais) variáveis aleatórias X, Y (facilmente generalizável para N) é $f_{X,Y}(x, y)$, tal que $f_{X,Y}(x, y)dx dy = P(x \leq X \leq x + dx, y \leq Y \leq y + dy)$. O par (X, Y) representa uma VA bidimensional. A densidade conjunta acumulada é $F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$.

A FDP *marginal*, por exemplo, da variável X é dada por

$$f_X(x) = \int_{-\infty}^{\infty} dy f_{X,Y}(x, y), \quad (5)$$

enquanto a densidade marginal acumulada é $F_X(x) = F_{X,Y}(x, \infty)$.

A FDP *condicionada* de X , dado um certo valor de $Y = y$, é

$$f_{X|Y}(x, y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \quad \text{com } f_Y > 0, \quad (6)$$

onde f_Y é a FDP marginal de Y .

Q. 3: Qual é a expressão para $f_{X|Y}$ quando X e Y são independentes.

Q. 4: Para as FDPs de Poisson e exponencial calcule $P(X > s + t | X > s)$ e interprete.

1.3 Momentos, função característica

O *momento de ordem n* da variável X é dado por:

$$\langle X^n \rangle = \int_{-\infty}^{\infty} dx x^n f_X(x). \quad (7)$$

O momento *centrado* de ordem n da variável X é:

$$\langle [X - \mu_X]^n \rangle = \int_{-\infty}^{\infty} dx [x - \mu_X]^n f_X(x), \quad (8)$$

onde $\mu_X \equiv \langle X \rangle$ é a *média* (aritmética) ou *valor esperado*. A *variância* ou *desvio quadrático* é o momento centrado de ordem 2, $\langle [X - \mu_X]^2 \rangle$, e o *desvio padrão* a sua raiz quadrada:

$$\sigma_X = \sqrt{\langle [X - \mu_X]^2 \rangle}. \quad (9)$$

Q. 5: Sendo α e β constantes, expresse a variância de $(\alpha X + \beta)$ em termos de σ_X^2 .

Q. 6: Calcule a variância de $Y = X_1 + X_2$, sendo X_1 e X_2 independentes. Generalize para a soma de $n > 2$ variáveis aleatórias independentes.

Q. 7: A variância é sempre uma quantidade finita? e a média? Dê exemplos.

A *função característica* (FC) da variável X é a transformada de Fourier da sua FDP

$$G_X(k) = \langle e^{ikx} \rangle = \int_{-\infty}^{\infty} dx e^{ikx} f_X(x). \quad (10)$$

G_X pode ser considerada uma função geratriz de momentos já que os coeficientes da expansão em série de Taylor estão relacionados com os momentos da FDP associada

$$G_X(k) = \sum_{m \geq 0} \frac{(ik)^m}{m!} \langle X^m \rangle, \quad (11)$$

que podem ser obtidos como sendo

$$\langle X^n \rangle = (-i)^n \left. \frac{d^n G_X}{dk^n} \right|_{k=0}. \quad (12)$$

Inversamente, conhecendo os momentos de uma distribuição, podemos conhecer a distribuição antitransformando (10). A FC também permite obter os *cumulantes* $\langle\langle X^m \rangle\rangle$ (combinações dos momentos), definidos a partir da série

$$\ln G_X(k) = \sum_{m \geq 1} \frac{(ik)^m}{m!} \langle\langle X^m \rangle\rangle. \quad (13)$$

Q. 8: Sendo α e β constantes, calcule a FC de $(\alpha X + \beta)$ em termos de G_X .

Q. 9: Expresse G_{X+Y} em termos de G_X e G_Y , sendo X e Y independentes. Generalize para N variáveis independentes.

Q. 10: Calcule a FC de uma variável com distribuição normal $N(0,1)$ e a partir dela os momentos e cumulantes. Calcule a FC da soma de duas variáveis com essa distribuição.

Q. 11: Calcule a FC de uma variável com distribuição lorentziana (2-gaussiana na tabela da Sec. 1.4) e a partir dela conclua sobre os momentos. Calcule a FC da soma de duas variáveis com essa distribuição.

1.4 Algumas distribuições de probabilidade unidimensionais usuais

denominação	distribuição/densidade	média	variância
binomial	$P(k) = \binom{n}{k} p^k (1-p)^{n-k}$ $k=0,1,\dots,n$ ($n=1,2,\dots$; $0 \leq p \leq 1$)	np	$np(1-p)$
de Poisson	$P(k) = e^{-\alpha} \alpha^k / k!$ $k=0,1,\dots$ ($\alpha > 0$)	α	α
geométrica	$P(k) = p(1-p)^{k-1}$ $k=1,2,\dots$ ($0 \leq p \leq 1$)	$1/p$	$(1-p)/p^2$
exponencial	$f(x) = \alpha e^{-\alpha x}$ $x \geq 0$ ($\alpha > 0$)	$1/\alpha$	$1/\alpha^2$
gaussiana ou normal, $N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
q -gaussiana	$f(x) \sim [1 + \beta(q-1)(x-\mu)^2]_+^{\frac{1}{1-q}}$ ($\beta > 0$)	calcule	calcule
lorentziana ou de Cauchy	$f(x) = \frac{\gamma/\pi}{(x-\mu)^2 + \gamma^2}$ ($\gamma > 0$)	—	∞
t de Student	$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \frac{1}{(1+x^2/n)^{\frac{n+1}{2}}}$ ($n=1,2,\dots$)	0 ($n > 1$) — ($n=1$)	∞ ($n \neq 2$) $\frac{n}{n-2}$ ($n > 2$)
χ -quadrado, χ_n^2	$f(x) = \frac{x^{n/2-1} e^{-x/2}}{2^{n/2}\Gamma(n/2)}$ $x \geq 0$ ($n=1,2,\dots$)	n	$2n$
de Lévy, L_γ	$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \cos kx e^{-\alpha k ^\gamma}$ ($0 < \alpha, 0 < \gamma < 2$)	0	∞

Observe que as distribuições gaussiana, de Lorentz e t de Student são casos particulares da q -gaussiana, com $q \rightarrow 1$, $q = 2$ e $q = (n+3)/(n+1)$, respectivamente.

1.5 Descorrelação e independência estatística

A *covariância* de duas variáveis aleatórias X e Y é

$$\text{cov}(X, Y) = \langle [X - \mu_X][Y - \mu_Y] \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy [x - \mu_X][y - \mu_Y] f_{XY}(x, y) \quad (14)$$

e o *coeficiente de correlação* é dado por

$$\rho(X, Y) = \frac{\langle [X - \mu_X][Y - \mu_Y] \rangle}{\sigma_X \sigma_Y}. \quad (15)$$

Q. 12: Provar que $|\rho(X, Y)| \leq 1$ e que a igualdade vale se existe dependência linear entre X e Y .

A independência estatística de duas variáveis X e Y pode ser expressa alternativamente como:

- (I) A FC conjunta $G_{XY}(k_1, k_2)$ pode ser fatorada.
- (II) Todos os momentos $\langle X^n Y^m \rangle$, com $n, m \geq 0$, podem ser fatorados.
- (III) Todos os cumulantes $\langle\langle X^n Y^m \rangle\rangle$, com $n, m \geq 1$ são nulos.

Por outro lado, duas variáveis X e Y são ditas *descorrelacionadas* se a sua covariância é nula (mais fraco que independência).

Q. 13: Procure um exemplo de variáveis descorrelacionadas porém dependentes.

Q. 14: Mostre que, para a distribuição gaussiana bidimensional, descorrelação implica independência. Analise a validade dessa afirmação no caso multidimensional.

Q. 15: Para a variável $X = (X_1, \dots, X_n)$ com distribuição gaussiana multidimensional de média nula, ache os momentos $\langle X_i X_j \dots \rangle$.

1.6 Lei dos grandes números

Uma propriedade importante é a *desigualdade de Chebyshev*¹⁾: se a variável aleatória X tem média μ e variância σ^2 :

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}, \quad \forall k > 0. \quad (16)$$

Para variáveis aleatórias independentes X_i ($i = 1, \dots, N$) identicamente distribuídas com média μ vale

$$\lim_{N \rightarrow \infty} P\left(\left|\frac{X_1 + X_2 + \dots + X_N}{N} - \mu\right| \geq \epsilon\right) = 0, \quad \forall \epsilon > 0. \quad (17)$$

Esta proposição é conhecida como *lei fraca dos grandes números* (relativa à convergência em probabilidade). No caso particular em que a distribuição possua variância σ^2 finita, a lei anterior é uma consequência trivial da desigualdade de Chebyshev, fazendo $k = \epsilon\sqrt{N}/\sigma$.

Uma outra expressão da lei fraca dos grandes números é relativa a ensaios de Bernoulli. Consideremos que, repetindo a experiência um número grande de vezes N (independentes), o evento A ocorre N_A vezes. Seja p_A a probabilidade do evento (a mesma em cada repetição), então dado $\epsilon > 0$, de (16), temos

$$P\left(\left|\frac{N_A}{N} - p_A\right| \geq \epsilon\right) \leq \frac{p_A(1-p_A)}{N\epsilon^2}, \quad (18)$$

que tende para zero no limite $N \rightarrow \infty$.

1.7 Teorema do limite central

Sejam as X_i ($i = 1, \dots, N$) variáveis aleatórias independentes com média μ_i e desvio padrão σ_i (finitos), segundo o *teorema do limite central* (TLC) a distribuição da variável

$$Z = \frac{\sum_{i=1}^N X_i - \sum_{i=1}^N \mu_i}{\sqrt{\sum_{i=1}^N \sigma_i^2}},$$

tende para uma distribuição normal $N(0,1)$ quando $N \rightarrow \infty$, se é satisfeita a condição de Lindeberg ¹⁾.

Aqui esboçaremos somente uma demonstração para o caso em que as variáveis da soma Z são identicamente distribuídas com média μ e desvio σ . Neste caso

$$G_Z(k) = e^{-\frac{ik\sqrt{N}\mu}{\sigma}} G_X^N\left(\frac{k}{\sqrt{N}\sigma}\right). \quad (19)$$

Tomando \ln de ambos os membros desta igualdade, desenvolvendo $\ln G_X$ em potências de k e levando em conta que N é grande, se chega a $G_Z(k) = e^{-\frac{k^2}{2}}$, que é a FC correspondente à distribuição $N(0,1)$.

Q. 16: Ache numericamente a distribuição da soma de N variáveis aleatórias uniformes no intervalo $[-1, 1]$, com $N = 2$ e 10 . Discuta o resultado numérico com relação ao TLC. O resultado das simulações é apresentado na Fig. 1.

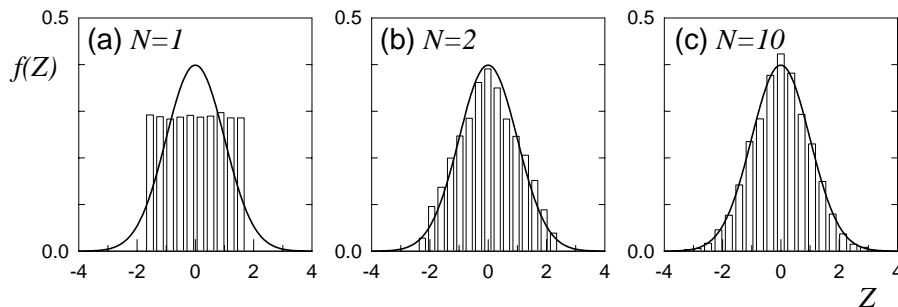


Figura 1: Resultados das simulações da Q. 16. Histogramas da variável normalizada Z calculada sobre 10^4 realizações para os valores de N indicados na figura.

1.8 Entropia

Um funcional importante da densidade de probabilidade é a entropia. Ela representa uma medida da desordem, da incerteza ou da desinformação. A forma entrópica mais usual, a de Shannon, é dada por

$$S(X) = - \int dx f_X(x) \ln f_X(x). \quad (20)$$

Porém muitas outras formas entrópicas podem ser encontradas na literatura (veja por exemplo ²⁾) que também verificam os requerimentos matemáticos de uma boa medida informacional, tais como as propriedades de positividade, concavidade definida, axiomas de Khinchin, etc.

Q. 17: Ache a distribuição definida num intervalo finito para a qual a entropia S é máxima?

Q. 18: Dadas duas variáveis independentes X e Y , ache a relação entre a entropia da distribuição conjunta e as entropias das distribuições marginais.

2 Processos estocásticos

Uma função aleatória $\mathcal{P}_X(t)$ é uma aplicação que atribui a cada valor do índice $t \in \mathbb{T}$ uma variável aleatória X_t . Um processo estocástico (PE) é uma função aleatória do tempo ^{3, 4}. O processo pode ser de tempo discreto ou contínuo dependendo da natureza do conjunto \mathbb{T} . Atribuindo-se um valor particular $x(t)$ à variável aleatória X_t (que alternativamente denotaremos $X(t)$), para cada t , temos uma *realização* do processo, uma função ordinária do tempo. Um PE $\mathcal{P}_X(t)$ também pode ser visto como o conjunto ou *ensemble* de todas as suas realizações $\{x(t), t \in \mathbb{T}\}$ ¹.

Um exemplo de PE é a seqüência de resultados ao se lançar muitas vezes um dado. Neste caso particular trata-se de um PE de tempo discreto em que as X_t são VA identicamente distribuídas e independentes. Outros exemplos são ilustrados na Fig. 2.

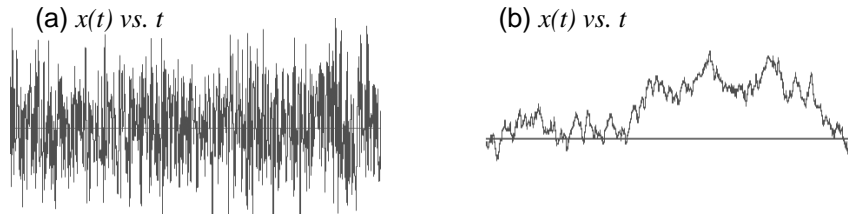


Figura 2: Realizações típicas de dois processos estocásticos: (a) ruído branco e (b) ruído browniano.

Antes de caracterizar diferentes tipos de PEs, vamos analisar um exemplo simples porém que permite ilustrar várias noções importantes que serão aprofundadas ao longo do curso, o *passeio aleatório*.

¹Veja também a interessante descrição alternativa de van Kampen ⁵).

2.1 “Intermezzo”: passeio aleatório unidimensional

Considere um caminhante inicialmente (i.e., a $k = 0$) posicionado na origem de coordenadas e que a cada unidade de tempo ($k = 1, 2, \dots$) anda sobre o eixo x com passos aleatórios independentes ξ_k . Suponhamos que as variáveis aleatórias ξ_k têm média nula e variâncias iguais σ_ξ^2 , então a variância da variável aleatória “posição após N passos”, $X_N = \sum_{k=1}^N \xi_k$, é

$$\sigma_{X_N}^2 = N \sigma_\xi^2. \quad (21)$$

Segundo o TLC a distribuição da variável X_N , para N suficientemente grande, é aproximadamente normal com média nula e variância $\sigma_{X_N}^2$.

Vejamos o caso em que os passos são de comprimento λ ou $-\lambda$ com probabilidades p e $q = 1 - p$, respectivamente. Neste caso, $P(X_N = n\lambda) = \binom{N}{m} p^m q^{N-m}$, com $m = (N + n)/2$, portanto $\langle X_N \rangle = (p - q) N\lambda$ e $\sigma_{X_N}^2 = 4pq N\lambda^2$. Consistentemente com a afirmação do parágrafo anterior, se N é grande, a distribuição binomial pode ser aproximada por uma gaussiana ⁶⁾ de tal forma que

$$P(n\lambda, N) \approx \frac{2\lambda}{\sqrt{2\pi\tilde{D}N}} e^{-\frac{(n-\tilde{v}N)^2}{2\tilde{D}N}}, \quad (22)$$

onde $P(n\lambda, N) \equiv P(X_N = n\lambda)$, $\tilde{v} \equiv (p - q)$ é a velocidade média e $\tilde{D} \equiv 4pq$ pode ser identificado com o coeficiente de difusão. Com efeito, é interessante notar que, por exemplo utilizando a relação de recorrência

$$P(n, N) = pP(n-1, N-1) + qP(n+1, N-1), \quad (23)$$

se chega a que a versão contínua de $P(n, N)$: $f(x, t) = P(n, N)/[2\lambda]$, onde $x = n\lambda$ e $t = N\tau$ (sendo λ , τ e $[p - q]$ apropriadamente pequenos ¹⁾), segue a equação diferencial

$$\frac{\partial f}{\partial t} = -v \frac{\partial f}{\partial x} + \frac{D}{2} \frac{\partial^2 f}{\partial x^2}, \quad (24)$$

onde $v = \tilde{v}\lambda/\tau$ e $D = \tilde{D}\lambda^2/\tau$. Esta equação diferencial é um caso particular da equação de Fokker-Planck que veremos mais adiante e corresponde a um processo difusivo normal com arrastamento. A FDP $f(x, t)$ também pode ser interpretada como a densidade de partículas (não-interagentes) na posição x ao tempo t . Assim, representa o deslocamento e dispersão de uma nuvem de partículas.

Q. 19: No caso particular em que $p = q = 1/2$, derive a equação de difusão a partir da relação de recorrência 23 .

Q. 20: Considere W caminhantes aleatórios independentes. Cada caminhante encontra-se inicialmente na origem de coordenadas e avança sobre o eixo x com passos unitários (± 1) equiprováveis em cada iteração. A partir de simulações computacionais:

- i) Calcule o desvio quadrático do deslocamento $\sigma^2 \equiv \sigma_{X_N}^2$ como função do tempo para $W = 1000$ e até tempo $N = 500$.
- ii) Para $W = 10000$ faça um histograma do número relativo de caminhantes em cada intervalo do eixo x para os tempos $N = 100$ e $N = 400$.
- iii) Compare os resultados obtidos com as previsões teóricas. Ver resultados na Fig. 3.

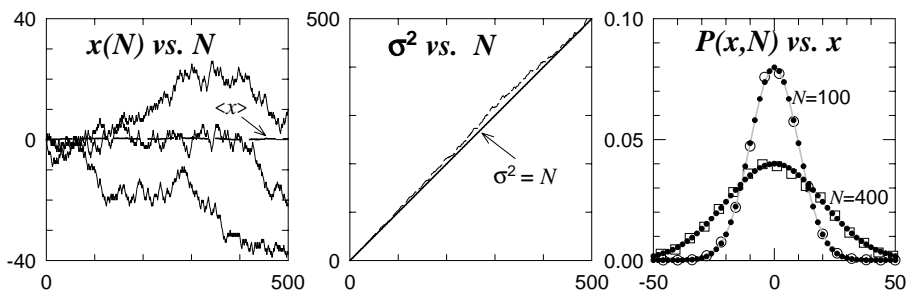


Figura 3: Resultados das simulações da Q. 20: (a) Diferentes realizações do processo, $x(t)$ (linhas contínua), e a média sobre $W = 10^3$ realizações (linha tracejada). (b) O desvio quadrático do deslocamento obtido numericamente para $W = 10^3$ (linha tracejada) pode ser descrito por $\sigma^2 = N$ (linha contínua). (c) Histogramas numéricos (símbolos) obtidos em $W = 10^4$ realizações e expressão analítica Eq. (22) (pontos) para os tempos $N = 100$ e $N = 400$.

2.2 Classes de processos estocásticos

As características que permitem distinguir entre PEs são a natureza do espaço de estados (espaço amostral de cada $X(t)$, com $t \in \mathbb{T}$), a natureza do conjunto de índices \mathbb{T} e as relações de dependência entre as variáveis aleatórias $X(t)$, que se derivam das distribuições conjuntas.

Conhecendo a hierarquia infinita de FDPs conjuntas correspondentes a n valores arbitrários da variável temporal $f_n(x_1, t_1; x_2, t_2; \dots; x_n, t_n) \equiv f_n(X(t_1) = x_1, \dots; X(t_n) = x_n)$, com $n = 1, 2, \dots$, o PE (seja discreto ou contínuo) fica

completamente definido ⁵⁾. As FDPs f_j da hierarquia com $1 \leq j < n$ são obtidas por integração de f_n . Dadas estas funções podemos calcular valores médios. Por exemplo, a função de autocorrelação de dois tempos é dada por

$$\langle X(t_1)X(t_2) \rangle = \int \int dx_1 dx_2 x_1 x_2 f_2(x_1, t_1; x_2, t_2). \quad (25)$$

Se as f_j não mudam ao substituir t_i por $t_i + \tau$ (com τ arbitrário), então trata-se de um *processo estacionário*.

Q. 21: Para processos estacionários, conclua sobre a dependência temporal de f_1 e f_2 .

O PE é dito estacionário em sentido amplo, se possui segundos momentos finitos e se $\text{cov}(X(t_i), X(t_i + \tau))$ depende somente de $\tau, \forall t_i \in \mathbb{T}$.

Q. 22: Para o caminhante aleatório da Q. 20, calcule $\text{cov}(X_{N_0}, X_{N_0+N})$. É esse um processo estacionário?

Uma possível classificação dos processos estocásticos é a seguinte ⁴⁾:

2.2.1 Processos puramente aleatórios:

A FDP condicional $f_{1|n-1}$, com $n > 1$, independe dos valores em instantes anteriores, ou seja

$$f_{1|n-1}(x_n, t_n | x_1, t_1; \dots; x_{n-1}, t_{n-1}) = f_1(x_n, t_n). \quad (26)$$

Q. 23: Para processos puramente aleatórios, ache uma expressão para f_n .

2.2.2 Processos markovianos:

A FDP condicional $f_{1|n-1}$, com $n > 1$, depende somente do valor no instante anterior, ou seja, se o valor presente é conhecido exatamente, o conhecimento

futuro não é alterado por informação adicional sobre o passado. Mais formalmente, para qualquer conjunto de tempos sucessivos $t_1 < t_2 < \dots < t_n$:

$$f_{1|n-1}(x_n, t_n | x_1, t_1; \dots; x_{n-1}, t_{n-1}) = f_{1|1}(x_n, t_n | x_{n-1}, t_{n-1}). \quad (27)$$

Um processo de Markov com espaço de estados discreto é uma *cadeia de Markov* (o processo de Poisson é uma cadeia de Markov de tempo contínuo). Um processo de Markov em que as realizações são funções contínuas é chamado *um processo de difusão* (o processo de Wiener é um caso típico).

Q. 24: Para processos markovianos, ache uma expressão para f_n .

2.2.3 Processos mais gerais:

Uma possível generalização dos PEs descritos acima corresponde ao caso em que as FDPs condicionais dependem de dois ou mais tempos.

2.3 Alguns processos particulares

2.3.1 Processo de Wiener:

Um processo $\mathcal{P}_X(t)$ de tempo contínuo (i.e, $\mathbb{T} = [0, \infty)$) é um processo de Wiener se para quaisquer $t_0 < t_1 < \dots < t_n$:

- os incrementos $X_{t_i} - X_{t_{i-1}}$ são mutuamente independentes (processo com incrementos independentes),
- a FDP do incremento $X_{t_i} - X_{t_{i-1}}$ depende somente de $t_i - t_{i-1}$ e
- a PDF do incremento $X_t - X_s$ é $N(0, \beta[t - s])$, com $\beta > 0$ para $t > s$. Em outros termos, o deslocamento $X_t - X_s$ pode ser visto como soma de um grande número de pequenos deslocamentos, tal que vale o TLC.

2.3.2 Processo de Poisson:

Seja um processo $\mathcal{P}_X(t)$ de tempo contínuo em que a variável aleatória X_t conta o número de vezes que ocorre um dado "evento" durante o intervalo $[0, t)$ e $p_n(t) = P(X_t = n)$, com $n = 0, 1, 2, \dots$. Um processo de Poisson é caracterizado pelas seguintes propriedades ¹⁾:

- as variáveis que representam o número de ocorrências em intervalos disjuntos são variáveis aleatórias independentes,
- se Y_t representa o número de ocorrências durante $[t_o, t_o + t)$, então para qualquer $t_o > 0$, X_t e Y_t têm a mesma pdf.
- $p_1(\Delta t) \simeq \lambda \Delta t$, onde λ é uma constante positiva, se $\Delta t > 0$ for suficientemente pequeno.
- $\sum_{k \geq 2} p_k(\Delta t) \simeq 0$, ou seja, a probabilidade de duas ou mais ocorrências durante um intervalo suficientemente pequeno é desprezível.
- $X_0 = 0$, ou seja, $p_0(0) = 1$.

Estas hipóteses permitem deduzir uma expressão para $p_n(t)$ chegando-se a $p_n(t) = e^{-\lambda t} (\lambda t)^n / n!$ (ver por exemplo ¹⁾).

2.3.3 Processo de Lévy:

Um processo $\mathcal{P}_X(t)$ com incrementos independentes se reduz a uma seqüência de variáveis aleatórias independentes $Z_0 = X_{t_0}$, $Z_i = X_{t_i} - X_{t_{i-1}}$, $i > 0$. Conhecendo as suas distribuições pode-se conhecer a distribuição conjunta de qualquer conjunto finito das $X_{t_i} = \sum_{i>0} Z_i$.

Se a distribuição de $X_{t+h} - X_t$ depende somente de h , se diz que o processo tem incrementos *estacionários*.

Um processo com incrementos independentes estacionários é denominado processo de Lévy.

Q. 25: Os processos de Wiener e de Poisson são processos de Lévy?

2.3.4 Martingalas:

Um PE $\mathcal{P}_X(t)$ (discreto ou contínuo) é uma martingala se $\langle X_t \rangle < \infty$ para todo t e se para quaisquer $t_1 < t_2 < \dots < t_{n+1}$ vale ⁷⁾:

$$\langle X_{t_{n+1}} | X_{t_1} = x_1, \dots, X_{t_n} = x_n \rangle = x_n < \infty.$$

Este é um modelo de jogos justos, em que a variável ao tempo t é X_t . A propriedade de martingala diz que a média ao tempo t_{n+1} , dado que se tem x_n ao tempo t_n , é igual a x_n independentemente de quais foram os valores passados.

3 Dinâmica estocástica: equação de Langevin

A dinâmica de um sistema estocástico pode ser descrita incorporando nas equações de movimento deterministas uma força flutuante (aleatória) apropriada, obtendo-se assim uma equação de Langevin. Muitos são os problemas que podem ser abordados a partir do tratamento estocástico da dinâmica, desde reações químicas até a dinâmica de populações. Um tratamento alternativo consiste na descrição probabilística através de uma equação para a distribuição de probabilidade dos estados do sistema. Veremos em detalhe estas duas descrições e exploraremos a conexão entre ambas. Antes, vamos analisar um caso concreto, o da partícula browniana.

3.1 Movimento browniano

Uma partícula no domínio browniano obedece a equação de movimento

$$\dot{v} = -\gamma v + \eta(t), \quad (28)$$

onde a força exercida pelas partículas do meio (dividida pela massa m da partícula), denominada força de Langevin, tem as seguintes propriedades:

- i. Consiste de um termo viscoso, que depende linearmente da velocidade v da partícula, e de um termo flutuante no tempo, independente de v .
- ii. A sua média é dada pelo termo viscoso de modo que $\langle \eta(t) \rangle = 0$, onde a média é tomada sobre um ensemble de sistemas.
- iii. Os choques experimentados pela partícula são descorrelacionados uns dos outros e ocorrem instantaneamente, tal que $\langle \eta(t) \eta(t') \rangle = \Gamma \delta(t - t')$, onde Γ é uma constante.

A equação (28) com estas propriedades representa uma *equação de Langevin* ^{4, 5}). A sua solução é

$$v(t) = v_0 e^{-\gamma t} + e^{-\gamma t} \int_0^t dt' e^{\gamma t'} \eta(t'), \quad (29)$$

onde v_0 é a velocidade inicial.

Da condição (ii), temos $\langle v \rangle = v_0 e^{-\gamma t}$, enquanto da condição (iii), obtemos

$$\langle v^2(t) \rangle = v_0^2 e^{-2\gamma t} + \frac{\Gamma}{2\gamma} (1 - e^{-2\gamma t}). \quad (30)$$

No equilíbrio térmico, considerando a lei de equipartição, a energia cinética média da partícula browniana se movimentando em 1D é dada por

$$\frac{1}{2}m\langle v^2 \rangle = \frac{1}{2}k_B T, \quad (31)$$

onde k_B é a constante de Boltzmann e T a temperatura. Então, para tempos $t \gg 1/\gamma$, de (30) e (31) resulta a relação $\Gamma = 2\gamma k_B T/m$, uma manifestação simples do *teorema de flutuação-dissipação* ⁶⁾.

Para o deslocamento, obtemos

$$x(t) = x_0 + v_0(1 - e^{-\gamma t})/\gamma + \int_0^t dt' (1 - e^{-\gamma(t-t')})\eta(t')/\gamma, \quad (32)$$

onde x_0 é a posição inicial. Da condição (ii) da força de Langevin, $\langle x \rangle = x_0 + v_0(1 - e^{-\gamma t})/\gamma$. Da condição (iii):

$$\langle [x(t) - \langle x \rangle]^2 \rangle = \frac{\Gamma}{\gamma^2} \left(t - \frac{2}{\gamma}(1 - e^{-\gamma t}) + \frac{1}{2\gamma}(1 - e^{-2\gamma t}) \right). \quad (33)$$

Q. 26: Interprete a expressão assintótica de $\langle (x - \langle x \rangle)^2 \rangle$ para tempos longos.

Para calcular momentos de ordem superior é necessário especificar as funções de correlação de mais de dois tempos. No caso particular em que o ruído é gaussiano (processo de Wiener, como suposto usualmente) as funções de correlação de mais de dois tempos ficam determinadas pelas correlações de dois tempos (ver Q. 15).

3.2 Outras equações de Langevin

Se a partícula em movimento unidimensional está sujeita adicionalmente a uma força externa $F(x)$, temos

$$\dot{v} = -\gamma v + F(x) + \eta(t), \quad (34)$$

onde o ruído possui as mesmas propriedades descritas anteriormente em conexão com a Eq. (28). Adicionalmente temos

$$\dot{x} = v. \quad (35)$$

Este sistema é um exemplo simples do caso mais geral em que as equações de movimento do sistema descrito pela variável aleatória $\mathbf{x} = (x_1, \dots, x_N)$ são da forma

$$\dot{x}_i = F_i(\mathbf{x}) + \eta_i(t), \quad (36)$$

tal que $\langle \eta_i \rangle = 0$ e $\langle \eta_i(t)\eta_j(t') \rangle = \Gamma_{ij}\delta(t-t'), \forall i, j$. Se as forças são lineares (ou seja $F_i = \sum_j F_{ij}x_j$), o processo é denominado de *Ornstein-Uhlenbeck*.

A equação de Langevin geral, no caso unidimensional, é da forma:

$$\dot{z} = F(z, t) + G(z, t)\eta(t), \quad (37)$$

onde $\eta(t)$ tem as propriedades estocásticas já descritas, mas note que a intensidade do ruído depende explicitamente do estado da partícula. Neste caso o ruído é denominado *multiplicativo*.

Q. 27: Considere a Eq. de Langevin $\dot{z} = z\eta(t)$, calcule os momentos de z como função do tempo sendo η um ruído gaussiano δ -correlacionado e com média nula.

3.3 Integral estocástica

Se $\eta(t)$ corresponde a uma seqüência aleatória de funções delta, então o termo de ruído provoca saltos em z de tal modo que o valor de z no instante do chute é indeterminado e portanto o valor de $G(z, t)$ nesse instante. Então a Eq. (37) está mal definida e é necessário introduzir uma especificação adicional que permita avaliar z no instante do salto. Esta dificuldade não ocorre quando o ruído é aditivo. Também não ocorreria se a correlação do ruído tivesse uma largura finita ϵ , mas neste caso o processo não seria markoviano.

Segundo a interpretação de Itô, deve ser considerado o valor antes do salto, de tal modo que a equação (37) pode ser escrita como

$$z(t + \Delta t) - z(t) = F(z(t), t)\Delta t + G(z(t), t) \int_t^{t+\Delta t} \eta(t') dt'. \quad (38)$$

Uma outra interpretação é a de Stratonovich, que considera o valor médio, em cujo caso obtém-se

$$z(t + \Delta t) - z(t) = F(z(t), t)\Delta t + G(\bar{z}, t) \int_t^{t+\Delta t} \eta(t') dt', \quad (39)$$

onde $\bar{z} = [z(t + \Delta t) + z(t)]/2$.

Representar graficamente uma realização de ruído branco δ -correlacionado é impossível. No movimento browniano (processo de Wiener), por exemplo, o

caminho $W(t) = \int_0^t \eta(t') dt'$ é contínuo porém não é diferenciável. Entretanto o seu incremento

$$w(\tau) = W(t + \tau) - W(t) = \int_t^{t+\tau} \eta(t') dt', \quad (40)$$

cuja média é nula e a sua autocorrelação $\langle w(\tau_1)w(\tau_2) \rangle = \min(\tau_1, \tau_2)$, é um novo processo de Wiener que resulta bem definido.

3.4 Derivação de FDPs a partir da equação de Langevin

Para achar distribuições de probabilidade, vamos discretizar as equações de Langevin e utilizar resultados conhecidos para a soma de VAs independentes 8).

Discretizando o tempo ($t = k\tau$, com $k \in \mathbb{Z}$), a equação de Langevin (28) pode ser escrita como 4)

$$v_{k+1} = \alpha v_k + \sqrt{\Gamma\tau} \xi_k, \quad (41)$$

onde $\alpha = (1 - \gamma\tau)$ e as VAs ξ_k têm as propriedades $\langle \xi_k \rangle = 0$ e $\langle \xi_k \xi_j \rangle = \delta_{kj}$. A velocidade resulta

$$v_k = v_0 \alpha^k + \sum_{s=0}^{k-1} w_s, \quad (42)$$

onde $w_s = \alpha^s \sqrt{\Gamma\tau} \xi_{k-s-1}$. Vamos supor que as VA ξ_k são independentes. Sendo $v_k - v_0 \alpha^k$ soma de VAs independentes com média nula e variância $\alpha^{2s} \Gamma\tau$, a sua distribuição para tempo longo segundo o TLC é gaussiana com média nula e variância $\tilde{\sigma}_k^2 = \sum_{s=0}^{k-1} \alpha^{2s} \Gamma\tau$.

No limite contínuo se chega a

$$f(v, t) = \frac{1}{\sqrt{2\pi\tilde{\sigma}(t)}} e^{-\frac{(v-c(t))^2}{2\tilde{\sigma}(t)^2}}, \quad (43)$$

onde $c(t) = v_0 e^{-\gamma t}$ e $\tilde{\sigma}^2(t) = \frac{\Gamma}{2\gamma}(1 - e^{-2\gamma t})$. Compare com os resultados da secção 3.1. Assintoticamente (no estado estacionário), $\tilde{\sigma}(t) \rightarrow \frac{\Gamma}{2\gamma} = \frac{k_B T}{m}$, obtendo-se a distribuição de Maxwell.

Para as posições, procedendo de modo semelhante temos

$$x_{k+1} = x_k + \tau v_k, \quad (44)$$

de onde resulta

$$x_k = x_0 + \tau \sum_{s=0}^{k-1} v_s = x_{k,0} + \sum_{s=1}^{k-1} u_{k,s}, \quad (45)$$

onde v_s é dado pela Eq. (42), $x_{k,0} = x_0 + v_0 \frac{1}{\gamma}(1 - \alpha^k)$ e $u_{k,s} = (1 - \alpha^s) \frac{1}{\gamma} \sqrt{\Gamma} \tau \xi_{k-s-1}$. Assim, no último membro temos de novo uma soma de VAs independentes. Considerando o caso particular $x_0 = v_0 = 0$, no limite contínuo se chega a

$$f(x, t) = \frac{1}{\sqrt{2\pi\sigma(t)}} e^{-\frac{x^2}{2\sigma^2(t)}}, \quad (46)$$

onde $\sigma(t)$ é o desvio quadrático médio da posição.

Q. 28: Faça esta derivação detalhadamente e mostre que para tempo longo $\sigma(t) \rightarrow Dt$, com D constante.

3.5 Evolução temporal dos momentos

Um outro possível caminho para obter FDPs consiste em determinar os momentos como função do tempo e, a partir deles, a função característica cuja antitransformada dá a FDP.

Vamos calcular os momentos como função do tempo para as posições da equação de Langevin com forças externas (34) no regime superamortecido. Neste caso a expressão discretizada para as posições é:

$$x_{k+1} = x_k + \tau F_k + \sqrt{\Gamma} \tau \xi_k, \quad (47)$$

com $F_k = F(x_k)$, $\langle \xi_k \rangle = 0$, $\langle \xi_k \xi_j \rangle = \delta_{kj}$ e onde consideramos $\gamma = 1$. Calculando x_{k+1}^n , desprezando termos de ordem superior em τ e tomando média obtemos:

$$\langle x_{k+1}^n \rangle = \langle x_k^n \rangle + n\tau \langle x_k^{n-1} f_k \rangle + \frac{1}{2} n(n-1) \tau \Gamma \langle x_k^{n-2} \rangle, \quad (48)$$

de onde resulta uma equação para a evolução temporal dos momentos:

$$\frac{d}{dt} \langle x^n \rangle = n \langle x^{n-1} F(x) \rangle + \frac{1}{2} n(n-1) \Gamma \langle x^{n-2} \rangle. \quad (49)$$

Q. 29: A partir da Eq. (49) ache os primeiros momentos para os casos em que (i) $F(x) = C$ e (ii) $F(x) = -Kx$, onde C e K são constantes. Relacione os resultados com outros anteriores.

Q. 30: A partir da Eq. (49) ache a FDP da posição, em ausência de força externa.

4 Equação de Fokker-Planck

Nesta seção vamos seguir um possível procedimento para derivar a equação de evolução da FDP da partícula browniana ⁸⁾. O procedimento será ilustrado com o caso de movimento browniano unidimensional dado pela Eq. (34), no regime superamortecido.

Seja a função característica associada à FDP $f(x_n)$:

$$G_n(k) = \langle e^{ikx_n} \rangle = \int dx_n e^{ikx_n} f(x_n). \quad (50)$$

Partindo da Eq. (47), usando a independência das ξ_n e desenvolvendo em potências de τ , temos

$$\begin{aligned} G_{n+1}(k) &= \langle e^{ikx_{n+1}} \rangle = \langle e^{ik[x_n + \tau F(x_n)]} \rangle \langle e^{ik\sqrt{\Gamma\tau}\xi_n} \rangle \\ &\simeq G_n(k) + \tau \left(ik \langle F(x_n) e^{ikx_n} \rangle - \frac{\Gamma}{2} k^2 \langle e^{ikx_n} \rangle \right) \end{aligned} \quad (51)$$

Escrevendo as médias como integrais, integrando por partes, igualando os integrandos, dividindo por τ e tomando o limite $\tau \rightarrow 0$, se chega a

$$\frac{\partial}{\partial t} f(x, t) = -\frac{\partial}{\partial x} [F(x)f(x, t)] + \frac{\Gamma}{2} \frac{\partial^2}{\partial x^2} f(x, t). \quad (52)$$

Este é um caso particular da *equação de Fokker-Planck* (FP) [o caso associado à equação de Langevin (34)]. Esta equação de FP particular é denominada *equação de Smoluchowski*.

A equação de FP geral é da forma ⁴⁾:

$$\frac{\partial}{\partial t} f(\mathbf{x}, t) = -\sum_{i=1}^N \frac{\partial}{\partial x_i} [D_i(\mathbf{x}, t)f(\mathbf{x}, t)] + \frac{1}{2} \sum_{i,j=1}^N \frac{\partial^2}{\partial x_i \partial x_j} [D_{ij}(\mathbf{x}, t)f(\mathbf{x}, t)]. \quad (53)$$

Q. 31: Utilizando um procedimento semelhante ao seguido para achar a Eq. (52), mostre que a equação de FP associada à equação de Langevin (36) é

$$\partial_t f(\mathbf{x}, t) = -\sum_{i=1}^N \partial_{x_i} [F_i(\mathbf{x})f(\mathbf{x}, t)] + \frac{1}{2} \sum_{i,j=1}^N \Gamma_{ij} \partial_{x_i}^2 f(\mathbf{x}, t). \quad (54)$$

Q. 32: Utilizando o resultado da questão anterior, mostre que para a equação de Langevin (34): $\dot{v} = -\gamma v + F(x) + \eta(t)$, a equação para a densidade $f(x, v, t)$ é a seguinte equação de FP, também denominada *equação de Kramers* unidimensional:

$$\partial_t f = -v \partial_x f - F(x) \partial_v f + \gamma \partial_v (vf) + (\Gamma/2) \partial_{vv}^2 f. \quad (55)$$

4.1 Algumas generalizações da equação de FP

A equação de FP não linear

$$\frac{\partial}{\partial t} f(x, t) = -\frac{\partial}{\partial x} [F(x)f(x, t)] + \frac{\Gamma}{2} \frac{\partial^2}{\partial x^2} [f(x, t)]^\nu \quad (56)$$

e a equação de FP com derivadas fraccionárias

$$\frac{\partial}{\partial t} f(x, t) = -\frac{\partial}{\partial x} [F(x)f(x, t)] + \frac{\Gamma}{2} \frac{\partial^{2\nu}}{\partial x^{2\nu}} f(x, t), \quad (57)$$

com $\nu > 0$, modelam a difusão anômala quando $\nu \neq 1$.

Outra generalização interessante introduz efeitos de memória:

$$\frac{\partial}{\partial t} f(x, t) = \int_{-\infty}^t \left(-\frac{\partial}{\partial x} F(x, t - \tau) + \frac{\partial^2}{\partial x^2} D(x, t - \tau) \right) f(x, \tau) d\tau. \quad (58)$$

4.2 Solução estacionária

No caso unidimensional, a equação de FP pode ser escrita como

$$\frac{\partial}{\partial t} f(x, t) = -\frac{\partial}{\partial x} J(x, t), \quad (59)$$

onde $J(x, t)$ é a corrente de probabilidade. Integrando ambos os membros da Eq. (59), e supondo que a variável x é definida no intervalo $[a, b]$, obtemos

$$\frac{d}{dt} \int_a^b dx f(x, t) = J(b, t) - J(a, t), \quad (60)$$

Para que a probabilidade se conserve, a corrente deve ser a mesma em ambos os extremos do intervalo em que a variável x está definida. No estado estacionário, a corrente deve ser constante. Se as condições de contorno são "refletoras" [$J(b, t) = J(a, t) = 0$], então $J(x, t) = 0 \forall x$.

No caso da equação de Smoluchowski (52), a corrente é dada por

$$J(x, t) = F(x)f(x, t) - \frac{\Gamma}{2} \frac{\partial}{\partial x} f(x, t) \quad (61)$$

e a solução estacionária resulta ser

$$f_s(x) \propto \exp(-2U(x)/\Gamma), \quad (62)$$

onde $U(x)$ é o potencial da força externa [$F(x) = -U'(x)$].

5 Equação mestra

5.1 Equação de Chapman-Kolmogorov

Na seção 2.2.2 vimos que um processo markoviano $\mathcal{P}_X(t)$ é tal que

$$f_{1|n-1}(x_n, t_n | x_1, t_1; \dots; x_{n-1}, t_{n-1}) = f_{1|1}(x_n, t_n | x_{n-1}, t_{n-1}). \quad (63)$$

Um PE markoviano fica completamente definido pelas funções f_1 e $f_{1|1}$. Toda a hierarquia pode ser reconstruída a partir destas funções, por exemplo

$$\begin{aligned} f_3(x_1, t_1; x_2, t_2; x_3, t_3) &= f_2(x_1, t_1; x_2, t_2) f_{1|2}(x_3, t_3 | x_1, t_1; x_2, t_2) \\ &= f_1(x_1, t_1) f_{1|1}(x_2, t_2 | x_1, t_1) f_{1|1}(x_3, t_3 | x_2, t_2), \end{aligned} \quad (64)$$

sendo $t_1 < t_2 < t_3$.

Integrando a Eq. (64) na variável x_2 , obtêm-se

$$f_3(x_1, t_1; x_3, t_3) = f_1(x_1, t_1) \int dx_2 f_{1|1}(x_2, t_2 | x_1, t_1) f_{1|1}(x_3, t_3 | x_2, t_2), \quad (65)$$

ou equivalentemente, dividindo ambos membros por $f_1(x_1, t_1)$:

$$f_{1|1}(x_3, t_3 | x_1, t_1) = \int dx_2 f_{1|1}(x_3, t_3 | x_2, t_2) f_{1|1}(x_2, t_2 | x_1, t_1), \quad (66)$$

denominada equação de *Chapman-Kolmogorov*.

Se adicionalmente o processo é estacionário, a densidade de um estado independe do tempo e a probabilidade de transição $f_{1|1}$ depende somente da diferença dos tempos, assim podemos escrever

$$f_{1|1}(x_2, t_2 | x_1, t_1) = T_\Delta(x_2 | x_1), \quad \text{com } \Delta = t_2 - t_1, \quad (67)$$

portanto a Eq. (66) fica

$$T_{\Delta+\Delta'}(x_3 | x_1) = \int dx_2 T_{\Delta'}(x_3 | x_2) T_\Delta(x_2 | x_1). \quad (68)$$

A seguir vamos obter uma representação diferencial no limite $\Delta \rightarrow 0$.

5.2 Derivação da equação mestra

Para Δ suficientemente pequeno, a probabilidade de transição tem a forma

$$T_{\Delta}(x_2|x_1) = [1 - a(x_1)\Delta]\delta(x_2 - x_1) + W(x_2|x_1)\Delta + \mathcal{O}(\Delta^2), \quad (69)$$

onde $W(x_2|x_1)$ é a probabilidade de transição do estado x_1 para o estado x_2 por unidade de tempo e

$$a(x_1) = \int dx_2 W(x_2|x_1).$$

Introduzindo (69) em (68):

$$T_{\Delta+\Delta'}(x_3|x_1) = [1 - a(x_3)\Delta']T_{\Delta}(x_3|x_1) + \int dx_2 W(x_3|x_2)T_{\Delta}(x_2|x_1)\Delta', \quad (70)$$

que, dividindo por Δ' e tomando o limite $\Delta' \rightarrow 0$, dá

$$\frac{\partial}{\partial \Delta} T_{\Delta}(x_3|x_1) = \int dx_2 [W(x_3|x_2)T_{\Delta}(x_2|x_1) - W(x_2|x_3)T_{\Delta}(x_3|x_1)]. \quad (71)$$

Esta equação é denominada *equação mestra* e pode ser escrita da forma mais habitual

$$\frac{\partial f(x,t)}{\partial t} = \int dx' (W(x|x')f(x',t) - W(x'|x)f(x,t)). \quad (72)$$

Representa uma equação de ganho/perda da probabilidade em cada estado. Quando o espaço de fases é discreto, a equação mestra tem a forma

$$\dot{p}_n(t) = \sum_{n'} (W_{nn'}p_{n'}(t) - W_{n'n}p_n(t)). \quad (73)$$

O primeiro termo representa o ganho devido a transições de outros estados ($n' \rightarrow n$), o segundo a perda devido às transições para outros estados ($n \rightarrow n'$).

No equilíbrio vale

$$\sum_{n'} W_{nn'}p_{n'}^e = \sum_{n'} W_{n'n}p_n^e. \quad (74)$$

Em determinadas situações (sistemas isolados) vale a condição de *balanço detalhado*:

$$W_{nn'}p_{n'}^e = W_{n'n}p_n^e. \quad (75)$$

Esta condição representa a reversibilidade microscópica.

5.3 Entropia e irreversibilidade

Consideremos a equação mestra discreta (73) e suponhamos que esta admite uma solução estacionária discreta com $p_n^e > 0$. Seja $f(x)$ uma função não-negativa ($f(x) > 0$) e convexa ($f''(x) \geq 0$) definida para $x \in [0, \infty)$. A quantidade

$$H(t) = \sum_n p_n^e f(x_n), \quad (76)$$

onde $x_n \equiv \frac{p_n(t)}{p_n^e}$, verifica $H(t) \geq 0$ e

$$\dot{H}(t) = \sum_{nn'} W_{nn'} p_{n'}^e (x_n f'(x_n) - x_{n'} f'(x_{n'})). \quad (77)$$

Por outro lado, para qualquer conjunto de números ϕ_n verifica-se a identidade

$$\sum_{nn'} W_{nn'} p_{n'}^e (\phi_n - \phi_{n'}) = 0. \quad (78)$$

Escolhendo $\phi_n = f(x_n) - x_n f'(x_n)$ e somando a identidade resultante no segundo membro de (77), tem-se

$$\dot{H}(t) = \sum_{nn'} W_{nn'} p_{n'}^e ([x_{n'} - x_n] f'(x_n) + f(x_n) - f(x_{n'})). \quad (79)$$

Sendo f convexa, o fator entre parênteses é negativo exceto se $x_n = x_{n'}$. Portanto $H(t)$ decresce monotonicamente. Sendo $H(t)$ uma quantidade não-negativa, então deve tender a um limite, em cujo caso (79) é nula. Qualquer função $f(x)$ não-negativa e convexa pode ser usada, porém $f(x) = x \ln x$ (portanto $H = \sum_n p_n \ln[p_n/p_n^e]$) tem propriedades adicionais tal como a extensividade ⁵⁾.

5.4 Derivação da equação de Fokker-Planck

A equação de FP pode ser derivada como uma aproximação da equação mestra ⁵⁾. A probabilidade de transição $W(x|x')$ pode ser expressa como função do estado inicial e do salto $s = x - x'$: $W(x|x') \equiv W(x'; s)$. Assim a equação (72) resulta

$$\begin{aligned} \frac{\partial f(x, t)}{\partial t} &= \int ds \left(W(x'; s) f(x', t) - W(x; -s) f(x, t) \right) \\ &= \int ds \left(W(x - s; s) f(x - s, t) - W(x; -s) f(x, t) \right). \end{aligned} \quad (80)$$

A aproximação consiste em supor primeiro que $W(x', s)$ é uma função com um pico estreito em s mas que varia suavemente em x' . Mais precisamente, existe $\delta > 0$ tal que

$$\begin{aligned} W(x'; s) &\simeq 0 && \text{para } |s| > \delta \\ W(x' + y; s) &\simeq W(x'; s) && \text{para } |y| < \delta. \end{aligned} \quad (81)$$

Uma segunda aproximação consiste em considerar que a solução $f(x, t)$ também varia lentamente com x como no caso anterior, de tal modo que o deslocamento de x para $x - s$ pode ser tratado por meio de uma expansão em série de Taylor até a segunda ordem. Assim se chega a

$$\frac{\partial f(x, t)}{\partial t} \simeq - \int s \frac{\partial}{\partial x} [W(x; s) f(x, t)] ds + \frac{1}{2} \int s^2 \frac{\partial^2}{\partial x^2} [W(x; s) f(x, t)] ds, \quad (82)$$

que também pode ser escrita como

$$\frac{\partial f(x, t)}{\partial t} = - \frac{\partial}{\partial x} [A_1(x) f(x, t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [A_2(x) f(x, t)], \quad (83)$$

onde

$$A_k(x) = \int s^k \frac{\partial}{\partial x} W(x; s) ds. \quad (84)$$

Se tivessem sido considerados todos os termos da expansão teríamos

$$\frac{\partial f(x, t)}{\partial t} = \sum_{k=1}^{\infty} \left\{ \frac{(-1)^k}{k!} \frac{\partial}{\partial x} [A_1(x) f(x, t)] + \frac{1}{2} \frac{\partial^k}{\partial x^k} [A_k(x) f(x, t)] \right\}, \quad (85)$$

que é a denominada expansão de *Kramers-Moyal*. A aproximação de Fokker-Planck supõe que os termos com $k > 2$ são desprezíveis. O truncamento é exato no caso de ruído gaussiano δ -correlacionado. Os coeficientes da expansão podem ser calculados a partir da equação de Langevin seguindo, por exemplo, o procedimento em ⁴⁾.

5.5 Equação diferencial de Chapman-Kolmogorov

Uma equação mais geral é a denominada *equação diferencial de Chapman-Kolmogorov* ³⁾, em que a evolução da probabilidade contém as variações temporais devido tanto à equação de Fokker-Planck, Eq. (53), como à equação mestra, Eq. (72). Ela deriva da equação de Chapman-Kolmogorov relaxando a hipótese de saltos pequenos (considerada na derivação da equação de FP na seção anterior) mas separando as contribuições devido a pequenos ($|s| < \delta$) e grandes saltos ($|s| > \delta$).

6 Análise de séries temporais

Uma série temporal é uma série de observações ou medidas obtidas em diferentes instantes de tempo e pode ser vista como uma realização de certo processo estocástico. Em diversas áreas, tais como medicina diagnóstica, meteorologia, finanças, etc., os sistemas de interesse são tipicamente estocásticos e somente se tem acesso a uma única ou poucas séries temporais de um dado processo. Devido à influência de múltiplos fatores envolvidos, as equações de movimento desses sistemas são em geral pobremente conhecidas. Portanto, uma questão central para essas aplicações é a elaboração de teorias que permitam prever a evolução futura a partir do conhecimento de uma ou poucas realizações. Em alguns casos em que os dados são muito ruidosos se procura achar ao menos leis universais, leis comuns a sistemas que podem ser muito diferentes, que permitem a caracterização de um dado fenômeno dentro de classes de universalidade. Estas leis podem ser dadas, por exemplo, por expoentes críticos, que dependem de poucos parâmetros tais como a dimensionalidade ou o número de parâmetros de ordem.

Uma questão importante para escolher o tipo de análise apropriado é decidir se os dados são estacionários ou não, ou seja se o sinal depende da origem de tempo da série. Em sistemas reais, os dados raramente são estacionários. Se as médias e outras propriedades estatísticas não variam significativamente com o tempo, ou seja não apresentam tendências no longo prazo, se diz que os dados são *quasi-estacionários* e na prática são tratados como sendo estacionários já que esta hipótese leva a simplificações teóricas. Assim, para lidar com dados não-estacionários é conveniente primeiro remover ou filtrar as componentes não-estacionárias. Também é importante detectar componentes periódicas. Em geral, um passo crucial é separar as componentes deterministas, das verdadeiramente estocásticas.

Uma das propriedades mais importantes que se procura determinar ao analisar séries temporais, é a denominada *persistência*, que se refere à memória ou correlação interna da série. Uma série é persistente se valores adjacentes estão correlacionados positivamente, como ocorre no movimento browniano. O ruído branco é o exemplo de série descorrelacionada. A série é dita *antipersistente* se valores adjacentes estão correlacionados inversamente, ou seja, se existe a tendência de valores grandes serem seguidos por valores pequenos.

Alguns dos efeitos mencionados (ciclos anuais, tendências, flutuações irregulares de curto prazo, etc.) podem ser observados na série real da Fig. 4. Discutiremos a seguir algumas técnicas de análise de séries temporais.

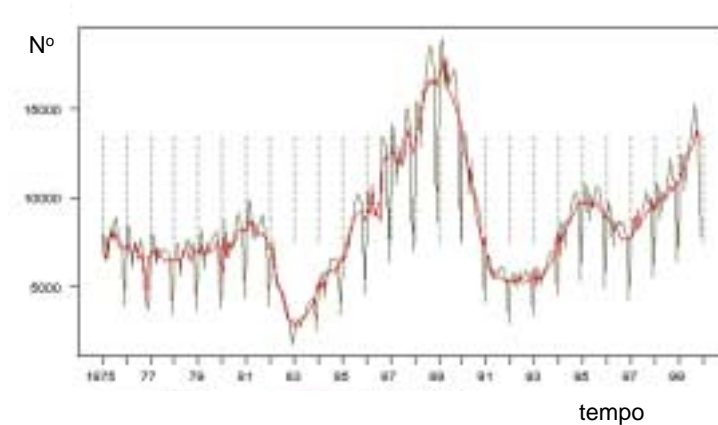


Figura 4: Série temporal do número de anúncios laborais mensais na Austrália ⁹⁾.

6.1 Densidade espectral e autocorrelação

Dado um PE $\mathcal{P}_X(t)$, se o processo é estacionário, a sua função de autocorrelação depende somente da diferença de tempos:

$$C(\tau) = \langle X(t + \tau)X^*(t) \rangle = \langle X(\tau)X^*(0) \rangle, \quad (86)$$

onde o asterisco denota o complexo conjugado. Se o processo é estacionário e *ergódico*, $C(\tau)$ independe do tempo e a média de ensemble pode ser substituída por uma média temporal sobre um intervalo de tempo suficientemente longo $2T$ de uma dada realização $x(t)$ do processo, assim

$$C(\tau) = \langle X(t + \tau)X^*(t) \rangle \simeq \frac{1}{2T} \int_{-T}^T dt' x(t' + \tau)x^*(t'). \quad (87)$$

A aplicabilidade destas hipóteses é crucial já que em geral se dispõe de uma única realização.

Por outro lado, dada a série $x(t)$, é útil conhecer sua decomposição em funções senoidais do tempo através da análise de Fourier. Considerando que $x(t)$ é nula fora do intervalo $-T \leq t \leq T$, a sua transformada de Fourier é

$$\tilde{x}(\omega) = \int_{-\infty}^{\infty} dt e^{-i\omega t} x(t) \quad (88)$$

[se $x(t)$ é real então $\tilde{x}(\omega) = \tilde{x}^*(-\omega)$] e a sua *densidade espectral*

$$S(\omega) \simeq \frac{1}{2T} |\tilde{x}(\omega)|^2. \quad (89)$$

Introduzindo a transformada (88) na definição (89), temos

$$\begin{aligned} S(\omega) &= \frac{1}{2T} \int_{-\infty}^{\infty} dt e^{i\omega t} x^*(t) \int_{-\infty}^{\infty} dt' e^{-i\omega t'} x(t') \\ &= \frac{1}{2T} \int_{-\infty}^{\infty} dt x^*(t) \int_{-\infty}^{\infty} d\tau e^{-i\omega\tau} x(t+\tau) \\ &= \int_{-\infty}^{\infty} d\tau e^{-i\omega\tau} \frac{1}{2T} \int_{-\infty}^{\infty} dt x(t+\tau)x^*(t) \\ &= \int_{-\infty}^{\infty} d\tau e^{-i\omega\tau} C(\tau). \end{aligned} \quad (90)$$

Este é um resultado importante, conhecido como *relação de Wiener-Khinchin*, que relaciona a densidade espectral $S(\omega)$ com a função de autocorrelação $C(\tau)$: a densidade espectral é a transformada de Fourier da autocorrelação e reciprocamente.

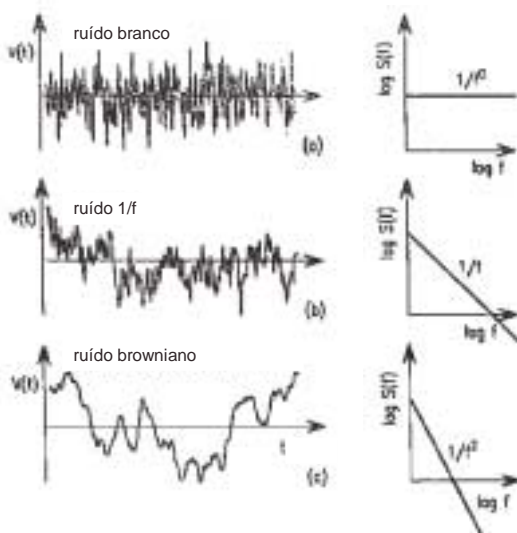


Figura 5: Densidade espectral $S(f)$ associada a realizações $v(t)$ de diferentes processos estocásticos 13).

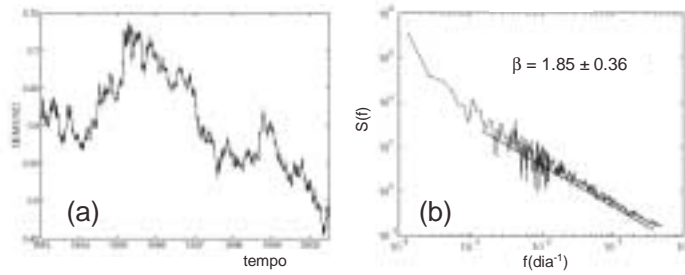


Figura 6: (a) Série temporal da taxa cambial marco-alemão/dólar no intervalo jan/1993-jun/2000. (b) A sua correspondente densidade espectral $S(f)$ ¹⁴.

Tipos comuns de ruído com as suas respectivas densidades espectrais são ilustrados na Fig. 5 em quanto na Fig. 6 é apresentado um caso real. Em todos estes casos, a densidade espectral segue uma lei de potência:

$$S(f) \propto f^{-\beta}. \quad (91)$$

Esta lei permite identificar uma classe de universalidade através do expoente β . Porém, a densidade espectral não sempre segue uma lei de potência como nesses casos. De todos modos, a transformada de Fourier informa se existem frequências mais abundantes que outras, em particular se existem ciclos, como pode ocorrer em meteorologia ou medicina. Portanto a transformada de Fourier dá sempre informação importante sobre a distribuição e correlações das amplitudes, frequências e fases das componentes do sinal ¹⁰.

Q. 33: Encontre a densidade espectral para o ruído δ -correlacionado $\eta(t)$ definido na seção 3.1. Como se modifica o resultado se a correlação tem uma largura finita?

Q. 34: Calcule a função de autocorrelação sabendo que a densidade espectral é da forma $S(f) \sim f^{-\beta}$.

A transformada em Fourier de $X(t)$ é um novo PE:

$$\tilde{X}(\omega) = \int_{-\infty}^{\infty} dt e^{-i\omega t} X(t). \quad (92)$$

A autocorrelação da variável transformada é:

$$\begin{aligned}
\langle \tilde{X}(\omega)\tilde{X}^*(\omega') \rangle &= \int_{-\infty}^{\infty} dt e^{-i\omega t} \int_{-\infty}^{\infty} dt' e^{i\omega' t'} \langle X(t)X^*(t') \rangle \\
&= \int_{-\infty}^{\infty} dt e^{-i\omega t} \int_{-\infty}^{\infty} dt' e^{i\omega' t'} \langle X(t-t')X^*(0) \rangle \\
&= \int_{-\infty}^{\infty} ds e^{-i(\omega-\omega')s} \int_{-\infty}^{\infty} ds' e^{-i(\omega+\omega')s/2} \langle X(s)X^*(0) \rangle \\
&= 2\pi\delta(\omega-\omega') \int_{-\infty}^{\infty} ds e^{-i\omega s} C(s) \\
&= 2\pi\delta(\omega-\omega')S(\omega).
\end{aligned}$$

Esta é outra das relações de Wiener-Khinchin e mostra que da estacionariedade de $X(t)$ resulta que, para frequências diferentes ω e ω' , $\tilde{X}(\omega)$ e $\tilde{X}^*(\omega')$ são descorrelacionadas.

6.2 Outros métodos de análise

6.2.1 Contagem de caixas: dimensão fractal

A *dimensão fractal* D é utilizada como uma medida da rugosidade do perfil. Permite quantificar a existência de padrões de auto-similaridade. Uma forma de medir D é pelo método de contagem de caixas:

$$D = \lim_{\epsilon \rightarrow 0} \frac{\ln N(\epsilon)}{\ln 1/\epsilon}, \quad (93)$$

onde $N(\epsilon)$ é o número mínimo de caixas de lado ϵ necessário para cobrir todo o conjunto de pontos. Porém, existem vários outros métodos, como o da integral de correlação de Grassberger-Procaccia (11).

6.2.2 Análise R/S: expoente de Hurst

Um outro expoente que permite medir o grau de rugosidade é o denominado *expoente de Hurst* H . Este expoente foi definido originalmente dentro da teoria do *intervalo rescalonado*, também denominada *análise R/S*. Esta teoria foi definida por Hurst para detectar a “persistência” ou memória de longo prazo em séries temporais.

O método consiste no seguinte procedimento: dada a série $z(t)$ no intervalo $0 \leq t \leq \tau$,

i. redefinir a variável como

$$y(t, \tau) = \sum_{i=1}^t [z(i) - \langle z \rangle_\tau], \quad (94)$$

onde a média é calculada no intervalo $[0, \tau]$. A média sobre o intervalo τ é subtraída para eliminar a tendência;

ii. determinar $R(\tau) = \max\{y(t, \tau)\} - \min\{y(t, \tau)\}$;

iii. calcular a variância nesse intervalo:

$$S^2(\tau) = \frac{1}{\tau} \sum_{i=1}^{\tau} [z(i) - \langle z \rangle_\tau]^2; \quad (95)$$

é esperado o comportamento $\frac{R(\tau)}{S(\tau)} \sim \tau^H$, de onde pode ser extraído o expoente característico H . Interpretando os $z(i)$ como incrementos espaciais de um passeio aleatório unidimensional, então a expressão (94) representa a posição do caminhante ao tempo t . No caso do movimento browniano unidimensional, o incremento da variável rescalonada no intervalo de comprimento τ , $R/S(\tau)$, é proporcional a $\tau^{1/2}$. Em geral o expoente pode ser diferente de 1: $R/S(\tau) \sim \tau^H$, com $0 \leq H \leq 1$, correspondendo a um tipo de processos denominado *movimento browniano fracionário* ¹³).

Pode ser mostrado que para perfis auto-afins $H = 2 - D$ e também que o expoente β da densidade espectral $[S(f)] \sim f^{-\beta}$ é $\beta = 2H + 1$. Se $1/2 < H \leq 1$, a série temporal é persistente, caracterizada por efeitos de memória de longo prazo.

Quanto menor H , maior é a dimensão fractal e portanto maior a rugosidade do sinal, segundo é ilustrado na Fig. (7).

6.2.3 DFA: expoente de Hausdorff

Um outro método de análise de séries é o denominado análise de flutuações sem tendências (DFA: "detrended fluctuation analysis" ¹²). Dada uma seqüência aleatória $y(n)$, são considerados N pontos em N/τ caixas, cada uma com τ pontos. Em cada caixa, se define o melhor ajuste linear $z(n) = an + b$. Considerando a função de flutuação $F(\tau)$

$$F^2(\tau) = \frac{1}{\tau} \sum_{n=(k-1)\tau+1}^{k\tau} |y(n) - z(n)|^2, \quad (96)$$

com $1 \leq k \leq N/\tau$. É esperado que

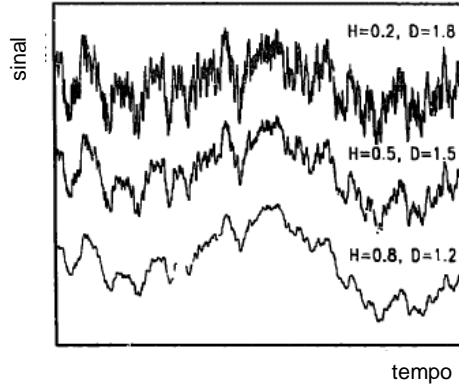


Figura 7: Valores de H e D para diferentes sinais ¹³⁾.

$$\langle F^2(\tau) \rangle^{1/2} \sim \tau^{H_a}, \quad (97)$$

onde a média é calculada sobre os N/τ intervalos. O expoente H_a , também denominado *expoente de Hausdorff*, coincide com H .

Referências

1. Feller, W., *Probability theory and its applications* (John Wiley, New York, 1950); Meyer, P.L., *Probabilidade, aplicações à estatística*, (Ao livro técnico, S.A., Rio de Janeiro, 1974).
2. Tsallis, C. J. Stat Phys. **52**, 479 (1988); Anteneodo, C. and Plastino, A.R. J. Phys. A **32**, 1089 (1999).
3. Gardiner, C. W., *Handbook of stochastic methods for physics, chemistry and the natural sciences* (Springer-Verlag, Berlin, 1983).
4. Risken, H., *The Fokker-Planck equation: methods of solution and applications* (Springer-Verlag, Berlin, 1984).
5. van Kampen, N.G., *Stochastic methods in physics and chemistry* (North-Holland Publ., Amsterdam, 1981).
6. Reif, F., *Fundamentals of statistical and thermal physics* (McGraw-Hill, Singapore, 1985).

7. Karlin, S. & Taylor, H.M., *A first course in stochastic processes* (Academic Press, New York, 1975).
8. Tomé, T. & de Oliveira, M.J., *Dinâmica estocástica e irreversibilidade* (Editora da Universidade de São Paulo, São Paulo, 2001)
9. Extraído da página do Australian Bureau of Statistics, <http://www.abs.gov.au/websitedbs/D3310114.nsf>
10. Jenkins G. M. & Watts, D. G., *Spectral Analysis and its Applications* (Holden-Day, San Francisco, 1968)
11. P. Grassberger & I. Procaccia, Phys. Rev. Lett. **50**, 346 (1983).
12. C.-K. Peng *et al.*, Phys. Rev. E **49**, 1685 (1994).
13. Voss, R.F. , Physica D **38** 362 (1989);
14. Ausloos, M., e-print, cond-mat/0103068 (2001).