UNIVERSIDADE ESTADUAL DE MARINGÁ
DEPARTAMENTO DE FÍSICA

Jean Hideaki Yamada Passos

Uma Revisão das Técnicas de Entropia e Complexidade de Permutação

Maringá, Outubro de 2018.

UNIVERSIDADE ESTADUAL DE MARINGÁ
DEPARTAMENTO DE FÍSICA

Jean Hideaki Yamada Passos

Uma Revisão das Técnicas de Entropia e Complexidade de Permutação

Trabalho de conclusão de curso apresentada ao Departamento de Física da Universidade Estadual de Maringá como requisito para obtenção do título de Bacharel em Física.

Orientador: Prof. Dr. Haroldo Valentin Ribeiro

Maringá, Outubro de 2018.

Agradecimentos

Agradeço a Deus acima de tudo.

Agradeço muito a meus pais, Ana e Leonardo por todo o suporte e compreensão em todos estes anos de estudo. Palavras não são suficientes para expressar todo o apoio que recebi de minha mãe.

Agradeço minha irmã Larissa pelo companheirismo e apoio, sem o qual eu não conseguiria progredir.

Agradeço imensamente ao professor Dr. Haroldo Valentin Ribeiro, pela disponibilidade e paciência sem tamanhos.

Agradeço meu amigo Sidney Cesar Sanches por me acompanhar pelo caminho das pedras por tantos anos.

E finalmente agradeço à UEM e ao Departamento de Física por fornecerem profissionais ímpares e essenciais à minha formação.

Resumo

Medidas de complexidade são onipresentes nos processos de caracterização de séries temporais, sejam elas de natureza física ou de qualquer outra. Trata-se de um aspecto particularmente importante para o estudo de sistemas complexos, nos quais, por muitas vezes, séries temporais sobre a dinâmica do sistema são a única informação disponível. Nesse contexto, apresentamos uma revisão das técnicas para análise de séries temporais chamadas entropia e complexidade de permutação. Essas medidas tem sido largamente utilizadas por conta de sua simplicidade, eficiência computacional e, principalmente, por poderem ser aplicadas a qualquer tipo de série temporal. Além dos aspectos formais da técnica, revisamos também algumas aplicações elementares a algumas séries temporais artificiais e empíricas.

Palavras-chave: Análise de dados. Física Estatística. Complexidade. Entropia. Séries temporais.

Abstract

Complexity measures are omnipresent in the process of time series characterization, whether they are of physical nature or any other. It is an aspect particularly crucial for the study of complex systems, where time series about the dynamics of a system are often the only available information. In this context, we aim to revise the techniques for time series analysis known as permutation entropy and permutation complexity. These measures are being largely used due to their simplicity, computational efficiency and, most importantly, because they can be applied to any time series. Beyond the formal aspects of the technique, we review some applications related to artificial time series and empirical ones.

Keywords: Data Analysis. Statistical Physics. Complexity. Entropy. Time Series.

Sumário

Introdução

O estudo de sistemas complexos depende muitas vezes de análises de sistemas que só podem ser descritos por séries temporais [1]. Séries temporais são extremamente comuns em nosso dia a dia: se seguirmos a variação de alguma quantidade com o tempo, estamos lidando com uma série temporal. Exemplos de séries temporais existem em grande quantidade e incluem desde os preços de ações na bolsa de valores, até o ícone que mostra o uso do processador de um computador com o decorrer do tempo. Outros exemplos incluem as séries das magnitudes de terremotos, temperatura, velocidade de vento, número de eventos numa guerra, som, eletrocardiogramas, etc. O que fazem séries temporais serem tão comuns e valiosas é que elas permitem enxergar não só um valor, mas uma história completa da dinâmica de um sistema. Com um dado valor e seu histórico, conseguimos reconhecer alterações nos padrões comuns de maneira relativamente fácil.

Para analisar séries temporais, diversos métodos e medidas podem ser empregados para caracterizar e diferenciar séries temporais de acordo com a sua complexidade. Isso ocorre porque o próprio conceito/definição de complexidade não está bem definido [2]. Exemplos de medidas de complexidade incluem complexidade algorítmica [3], entropias [4], entropias relativas [5], dimensões fractais [6] e expoentes de Lyapunov [7]. Essas medidas de complexidade são principalmente utilizadas para distinguir séries temporais periódicas de séries temporais caóticas. Entretanto, a maioria dessas medidas não pode ser aplicada a qualquer tipo de série temporal. Além disso, muitas delas dependem de algoritmos muito específicos que, muitas vezes, são bastante sensíveis a parâmetros de ajuste [8]. Uma consequência direta dessa característica é que surgem grandes dificuldades para reproduzir muitas dessas análises sem o conhecimento de pormenores dos métodos.

Devido a essas dificuldades, Bandt e Pompe [8] propuseram uma nova medida de complexidade chamada entropia de permutação. Essa medida de complexidade, considerada uma medida natural para a complexidade da série, contorna os problemas anteriores, podendo

ser aplicada a qualquer série temporal e tendo a vantagem de ser rápida do ponto de vista computacional. Outra característica positiva é a simplicidade dessa medida, tornando fácil a comparação entre sistemas diferentes, uma vez que elimina-se a necessidade de parâmetros de ajuste. Como veremos, a medida de Bandt e Pompe é definida como a entropia de Shannon das probabilidades associadas aos padrões ordinais que ocorrem numa série temporal.

A entropia de permutação, representa um método rápido e eficiente para descrever séries temporais. Porém, ela em si pode não ser suficiente em muitos contextos. Um dos principais empecilhos que a entropia de permutação encontra é o fato dela não conseguir distinguir entre séries caóticas e séries estocásticas. Motivados por essa limitação, Rosso et al. [9] propuseram uma extensão para a medida de Bandt e Pompe baseada no trabalho de Lopez-Ruiz et al. [10], chamada complexidade de permutação. De modo a diferenciar sistemas diferentes caracterizados pelo mesmo valor de entropia de permutação, Rosso et al. propõem multiplicar a entropia por uma quantidade chamada desequilíbrio. O desequilíbrio representa uma métrica entre a distribuição dos padrões ordinais da série e a distribuição equiprovável. Usando a entropia e complexidade, Rosso et al. construíram um diagrama no qual temos a entropia no eixo x e a complexidade no eixo y, o chamado plano de complexidade-entropia. Como esse plano, Rosso et al. mostraram que é possível distinguir entre séries temporais, as quais não podiam ser diferenciadas usando apenas os valores da entropia de permutação.

Com esse trabalho de conclusão de curso, pretendemos fazer um revisão dessas duas medidas de complexidade, apresentando seus aspectos formais e revisando algumas aplicações numéricas em séries artificiais e empíricas.

CAPÍTULO 1

Entropia de permutação

1.1 A entropia de permutação

A entropia de permutação é uma ferramenta poderosa para a análise de séries temporais. Ela tem sido utilizada cada vez mais, devido a sua velocidade e robustez, além de outras características positivas. Proposta por Christoph Bandt e Bernd Pompe [8] em 2002, os autores a propuseram como uma medida "natural" da complexidade de séries temporais, como alternativa às diversas outras medidas que já existem, apenas para citar algumas como complexidade algorítmica [3], entropias [4], entropias relativas [5], dimensões fractais [6] e expoentes de Lyapunov [7].

A entropia de Shannon é uma das mais importantes entropias que estão relacionadas à quantidade informacional, porém a maior dificuldade na descrição de um sistema é o cálculo das probabilidades relativas deste sistema [11]. Com o método da entropia de permutação, temos um método simples para realizar esta tarefa.

Outra vantagem da entropia de permutação é a sua capacidade de ser aplicada em séries temporais reais, devido à sua capacidade de analisar séries com ruídos [8,11]. Outras medidas de complexidade são eficientes apenas em aplicações à séries temporais simuladas, sem ruídos, falhando em medidas com ruídos. Para aplicar outras medidas em séries reais, é necessário muitas vezes a aplicação de um processamento de dados cuidadoso, de modo a remover o ruído, e com isso, é introduzida uma maior dificuldade na reprodução desses cálculos sem uma descrição completa de como o processamento dos dados foi feito. Ainda mais, o processo de remoção de ruídos em si acaba eliminando certas informações contidas na série temporal.

A essência da entropia de permutação é separar a série temporal em partições de tamanho

d, chamadas de "*embedding dimension*[1]". Separadas as partições d, fazemos um reordenamento de seus elementos em ordem crescente, de modo a gerar um padrão de permutação ($\pi(012)$ e assim por diante). Assim, a série temporal é separada em permutações, e cada permutação tem uma probabilidade de ocorrer na série, permitindo assim sua contagem e o cálculo de sua probabilidade, comparado com o número de partições totais que a série possui. O próprio termo "entropia de permutação" é derivado do método de simbolização e re-organização dos elementos proposto para obter a distribuição de probabilidade dos estados acessíveis ao sistema em questão.

Esse método é simples e eficiente pois ele leva em conta os valores vizinhos na criação de estados possíveis. Isso é de extrema importância, pois com esse método consegue-se englobar as informações importantes que estão codificadas na série temporal, um aspecto que geralmente não é levado em conta [11].

Agora vamos analisar a entropia de permutação de Bandt e Pompe [8] formalmente, e então aplicá-la a um exemplo simples para clarificar melhor como ela é aplicada em séries temporais reais.

Considere uma série temporal arbitrária composta de n elementos e representada por:

$$\mathcal{X} = \{x_1, x_2, \ldots, x_n\} = \{x_t\}_{t=1,2,\ldots,n}. \tag{1.1}$$

Dessa série temporal, construímos partições $\vec{s}$ de tamanho d, onde d é a *embedding dimension* escolhida:

$$(\vec{s}) \mapsto (x_{s-(d-1)}, x_{s-(d-2)}, \ldots, x_{s-1}, x_s), \tag{1.2}$$

com $\vec{s} = d, d+1, \ldots, n$. Por exemplo, se a nossa série temporal tem 7 elementos ($n = 7$) e escolhemos $d = 2$, temos que nossas partições serão as seguintes:

$$\begin{aligned} \vec{s} &= \vec{2}, \vec{3}, \vec{4}, \vec{5}, \vec{6}, \vec{7} \\ \text{para } \vec{2} &= (x_{2-(2-1)}, x_{2-(2-2)}) = (x_1, x_2) \\ \text{para } \vec{3} &= (x_{3-(2-1)}, x_{3-(2-2)}) = (x_2, x_3) \\ &\vdots \\ \text{para } \vec{7} &= (x_{7-(2-1)}, x_{7-(2-2)}) = (x_6, x_7) \end{aligned} \tag{1.3}$$

Separadas as partições, seus elementos internos recebem um padrão ordinal definido como a permutação de $\pi = (r_0 r_1 \ldots r_{d-1})$ dos símbolos $(0, 1, \ldots, (d-1))$, definidos pelo ordenamento

[1] A tradução seria algo como "dimensão de incorporação".

$$x_{s+r_{d-1}} \leq x_{s+r_{d-2}} \leq x_{s+r_1} \leq \ldots \leq x_{s-r_0} \tag{1.4}$$

Isso significa assimilar um padrão ao vetor, reodenar seus valores internos em ordem ascendente e então, o padrão permutado será o nosso padrão de permutação. Por exemplo, no caso em que $d = 2$, o vetor $\vec{2} = (x_1, x_2)$ tem um padrão ordinal 0 assimilado ao elemento x_1, e 1 assimilado ao elemento x_2. Reorganizando-os em ordem crescente, se $x_1 \leq x_2$, o padrão (01) se mantém, mas caso $x_2 \leq x_1$, o padrão (10) é criado.

Repetindo esse processo para todos os $(n - d + 1)$ vetores, obtemos todas as permutações contidas na série temporal. Como temos uma partição de tamanho d, temos que o número total de permutações possíveis é $d!$, calculamos a frequência com que cada uma dessas permutações aparece na nossa série temporal

$$p(\pi_i) = \frac{\#\{s | s \leq (n - d + 1);\ (\vec{s}) \text{ do tipo } \pi_i\}}{n - d + 1}, \tag{1.5}$$

em que o símbolo # representa o número de ocorrências de uma permutação e "$(\vec{s})$ do tipo π_i" significa os vetores $\vec{s}$ ordenados em uma permutação do tipo π_i. Em outras palavras, estamos dividindo o número total de vezes que uma permutação apareceu pelo número total de vetores. Com isso, temos o conjunto das probabilidades $P = \{p(\pi_i)\}_{i=1,2,\ldots,d!}$.

Usando esse conjunto de probabilidades (também denominado de distribuição dos padrões ordinais), calculamos a entropia de permutação de ordem d associada à série temporal $\mathcal{X}$ isto é,

$$S[P] = -\sum_{i=1}^{d!} p(\pi_i) \log p(\pi_i). \tag{1.6}$$

Vale notar que essa quantidade é a entropia de Shannon associada à distribuição P.

Como deve estar claro, o único parâmetro usado para determinar a entropia de permutação é d, a *embedding dimension*. Esse parâmetro é de grande importância, uma vez que d determina a quantidade de informação contida em cada vetor, além de $d!$ ser o número de estados acessíveis do sistema, formando o espaço amostral. A única condição a ser obedecida é $d! \ll n$ para que o cálculo de P seja confiável. Em seu trabalho, Bandt e Pompe recomendam que $d = (3, 4, \ldots, 7)$ para a aplicação da entropia de permutação em séries empíricas [8].

Analisando a equação 1.6 notamos que a entropia $S[P]$ varia entre $0 \leq S[P] \leq \log d!$, onde $\log d!$ é a entropia máxima, quando todos os estados são equiprováveis e o limite inferior $S[P] = 0$ acontece quando a série é regular. Com isso podemos definir a entropia de permutação normalizada

$$H[P] = \frac{S[P]}{S_{max}} = \frac{S[P]}{\log d!}. \tag{1.7}$$

Onde $S_{max} = \log d!$ é a distribuição que maximiza a entropia. Essa distribuição pode ser calculada via método dos multiplicadores de Lagrange [12] (conforme descrito no Apêndice A).

Com a introdução da normalização, limitamos o valor da entropia ao intervalo $0 \leq H[P] \leq 1$ onde novamente, $H[P] = 0$ para uma série regular e $H[P] = 1$ para uma série em que todas as $d!$ permutações possíveis são equiprováveis.

Agora com a medida formalmente definida, usaremos um exemplo simples para elucidá-la. Considere uma série $\mathcal{X}$ com $n = 7$ termos:

$$\mathcal{X} = \{10, 3, 6, 9, 2, 5, 8\}.$$

De posse da nossa série, o primeiro passo é escolher um tamanho apropriado da *embedding dimesion*. Escolhemos $d = 2$ pois nosso exemplo tem um número de termos n pequeno. O número total de vetores são $(n - d + 1) = 6$, que são representados por:

$$\begin{aligned}
(\vec{2}) &= (10, 3), \\
(\vec{3}) &= (3, 6), \\
(\vec{4}) &= (6, 9), \\
(\vec{5}) &= (9, 2), \\
(\vec{6}) &= (2, 5), \\
(\vec{7}) &= (5, 8).
\end{aligned}$$

O próximo passo é assinalá-los aos símbolos que gerarão a permutação π quando seus valores forem ordenados em ordem crescente:

$$\begin{aligned}
(\vec{2}) = (10, 3) &\rightarrow (a_1 < a_0) \rightarrow \pi(10) \\
(\vec{3}) = (3, 6) &\rightarrow (a_0 < a_1) \rightarrow \pi(01) \\
(\vec{4}) = (6, 9) &\rightarrow (a_0 < a_1) \rightarrow \pi(01) \\
(\vec{5}) = (9, 2) &\rightarrow (a_1 < a_0) \rightarrow \pi(10) \\
(\vec{6}) = (2, 5) &\rightarrow (a_0 < a_1) \rightarrow \pi(01) \\
(\vec{7}) = (5, 8) &\rightarrow (a_0 < a_1) \rightarrow \pi(01).
\end{aligned}$$

Assim, com o auxílio da tabela abaixo podemos contar o número que cada permutação aparece em nossa série temporal:

permutação “01”	permutação “10”
$(\vec{3}) = (3, 6)$	$(\vec{2}) = (9, 2)$
$(\vec{4}) = (6, 9)$	$(\vec{5}) = (5, 8)$
$(\vec{6}) = (2, 5)$	
$(\vec{7}) = (5, 8)$	

E as respectivas probabilidades $p(\pi)$ podem ser calculadas usando a equação 1.5:

$$p(\text{“01”}) = 4/6 = 0,66$$
$$p(\text{“10”}) = 2/6 = 0,33$$

Finalmente, definimos a entropia de permutação para essa série usando a equação 1.6, ou seja, calculamos a entropia de permutação[2] de ordem $d = 2$ como

$$S = -\frac{4}{6}\log\left(\frac{4}{6}\right) - \frac{2}{6}\log\left(\frac{2}{6}\right) \approx 0{,}918. \tag{1.8}$$

Utilizando a série anterior, podemos fazer outro exemplo tomando partições de tamanho $d = 3$. Para $d = 3$, temos $(n - d + 1) = 5$ vetores e $3! = 6$ possíveis permutações possíveis:

$$(\vec{3}) = (10, 3, 6),$$
$$(\vec{4}) = (3, 6, 9),$$
$$(\vec{5}) = (6, 9, 2),$$
$$(\vec{6}) = (9, 2, 5),$$
$$(\vec{7}) = (2, 5, 8),$$

Novamente precisamos analisar as permutações de cada partição:

$$(\vec{3}) = (10, 3, 6) \rightarrow (a_1 < a_2 < a_0) \rightarrow \pi(120)$$
$$(\vec{4}) = (3, 6, 9) \rightarrow (a_0 < a_1 < a_2) \rightarrow \pi(012)$$
$$(\vec{5}) = (6, 9, 2) \rightarrow (a_2 < a_0 < a_1) \rightarrow \pi(201)$$
$$(\vec{6}) = (9, 2, 5) \rightarrow (a_1 < a_2 < a_0) \rightarrow \pi(120)$$
$$(\vec{7}) = (2, 5, 8) \rightarrow (a_0 < a_1 < a_2) \rightarrow \pi(012).$$

Então, contando o número de vezes que cada permutação ocorre e calculando suas probabilidades:

$$p(\text{“012”}) = 2/5 = 0,4,$$
$$p(\text{“021”}) = 0,$$
$$p(\text{“102”}) = 0,$$
$$p(\text{“120”}) = 2/5 = 0,4,$$
$$p(\text{“201”}) = 1/5 = 0,2,$$
$$p(\text{“210”}) = 0,$$

[2]Nestes exemplos foram utilizados os logaritmos na base 2 e a constante $k = 1$, como o de costume no cálculo da entropia de Shannon: $S = -k\sum_{i=1}^{m} p_i \log p_i$.. No resto deste trabalho a entropia foi calculada na base e. Não há perda de generalidade devido ao fato de usarmos a entropia normalizada como na equação 1.7.

E então a entropia de permutação para $d = 3$ pode ser calculada usando a equação 1.6:

$$S = -2\left(\frac{2}{5}\right)\log\left(\frac{2}{5}\right) - \frac{1}{5}\log\left(\frac{1}{5}\right) \approx 1{,}522. \tag{1.9}$$

No caso de dois valores serem iguais, os elementos são organizados na ordem em que apareceram, por exemplo, se um vetor $\vec{s} = (1, 4, 1)$ estivesse presente, teríamos a permutação $\pi(021)$. Note também que para $d = 3$, das 6 possíveis permutações, apenas 3 aparecem em nosso exemplo. Esse é um fato de extrema importância, pois as probabilidades diferentes de cada vetor pode revelar informações importantes sobre o sistema. Quando uma permutação não aparece na série temporal, temos padrões proibidos. Essa dinâmica dos padrões proibidos e as diferentes probabilidades são o que permitem a entropia de permutação em analisar informações "escondidas" na série.

Temos duas razões principais para termos padrões proibidos na série:

- As séries temporais são compreendidas de um número finito de valores, consequentemente, um número finito de vetores;
- A dinâmica natural gerando a série temporal é o que limita o aparecimento de certos padrões.

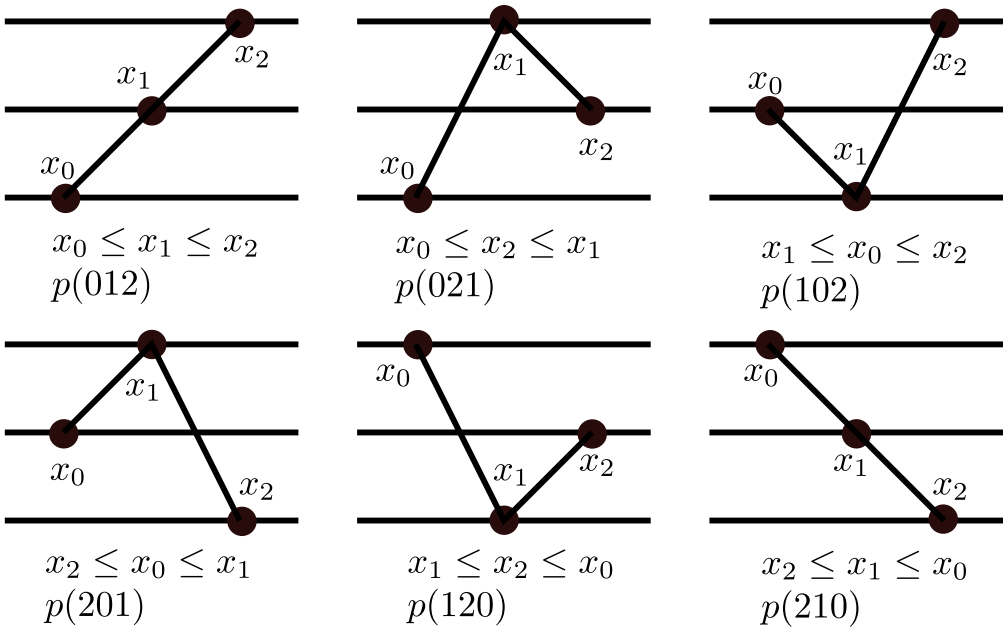

Figura 1.1: Representação dos vetores e suas respectivas permutações para $d = 3$.

Como pode ser visto, o procedimento para o cálculo dessas probabilidades consiste de um simples reordenamento local dos elementos, procedimento que pode ser feito de maneira muito simples e rápida do ponto de vista computacional, levando a robustez e velocidade desse procedimento.

Um exemplo feito por Bandt e Pompe para demonstrar a eficiência da aplicação da entropia de permutação em séries de dados reais, os autores analisaram uma série temporal gerada a partir das amplitudes do áudio de uma gravação[8]. Essa gravação, de aproximadamente 4 segundos, contém a pronúncia em inglês da frase *"permutation entropy measures complexity"* (entropia de permutação mede complexidade) por uma voz masculina em voz modal.

Como comparação, também foi usada uma técnica chamada ZCR ou *zero-crossing rate* [13]. A ZCR é uma medida de reconhecimento de fala bem difundida e é definida como a taxa com que o sinal passa pelo zero em um certo intervalo de tempo. Tanto a entropia de permutação H quanto a ZCR são usadas aqui para tentar identificar trechos na gravação que contém voz ativa. No caso da ZCR, o valor é perto de 0 quando há voz ativa, por volta de 0,5 quando há consoantes fricativas ou interrupções sonoras e um valor maior que 0,5 em chiados de fundo. A entropia de permutação, por sua vez, foi calculada utilizando $d = 3, 4, 5$, devido ao pequeno número de amostras e, em trechos de silêncio, a entropia normalizada $S[P]/log(d!)$ assume valores próximos de 1 e diminui em trechos de fala. A série temporal das amplitudes sonoras para o áudio pode ser vista na figura 1.2(a), junto com o resultado da entropia de permutação H em (b) e a ZCR em (c).

Analisando o resultado de H, percebe-se que a entropia de permutação consegue diferenciar com melhor eficiência o começo e o fim da amostra sonora, nas quais não há voz falada, apenas um ruído de fundo oriundo da rede elétrica, enquanto que o método ZCR apresenta um valor baixo devido a esse ruído. No trecho inicial de silêncio e nas pausas durante a fala, H assume valores próximos de 1, e o valor diminui nos trechos nos quais há voz. O "u" em "permutation" é melhor detectado por H e a transição em 1,6s, na qual a entropia de permutação registra voz, mas a ZCR não consegue identificar com clareza essa transição entre as sílabas. Assim, a entropia de permutação mostra-se um método melhor do que a ZCR, conseguindo quase que perfeitamente diferenciar as pausas e a fala, enquanto que a ZCR falha em vários pontos. No geral, a entropia H e a ZCR são similares, mas H consegue caracterizar melhor as sutilezas na fala.

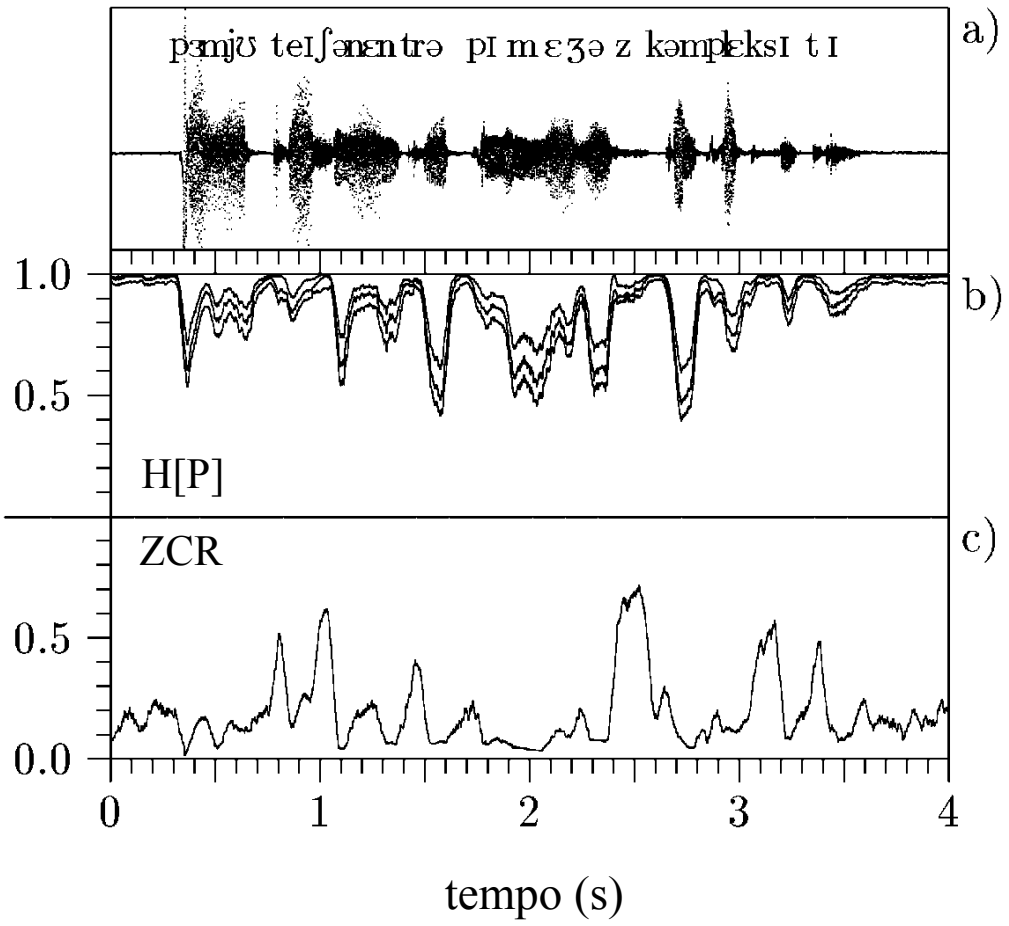

Figura 1.2: Análise da pronúncia da frase "entropia de permutação mede complexidade". (a) Representação da série temporal das amplitudes sonoras da gravação. Acima, a transcrição fonética da fala. (b) Entropia de permutação normalizada $H[P]$ para $d = 2, 3, 4$ (as linhas de cima para baixo). (c) Valor da ZCR para o mesmo trecho de áudio. Figura adaptada da referência [8]

Desse modo, a entropia de permutação é uma medida de complexidade apropriada para a caracterização de séries temporais de natureza arbitrária, com vários pontos positivos, entre os quais podemos listar [11]:

- Rapidez computacional;
- Robustez;
- Fácil reprodução;
- Apenas um parâmetro a ser ajustado (d);
- Eficiência na presença de ruídos;
- Aplicação em séries temporais arbitrárias.

Naturalmente a técnica não é perfeita para tudo e entre as poucas insuficiências/dificuldades da entropia de permutação podemos listar:

- Distinção entre séries estocásticas e caóticas;
- Sistemas de complexidades diferentes com a mesma entropia.

Essas insuficiências serão abordadas no próximo capítulo, com a introdução da complexidade de permutação.

CAPÍTULO 2

Complexidade Estatística de Permutação

Neste capítulo, estudaremos a complexidade de permutação, uma medida adicional a entropia de permutação, que nos permite refinar ainda mais o método. A complexidade é definida como a entropia de permutação multiplicada por uma quantidade nomeada desequilíbrio, que é definida pela divergência de Jensen-Shannon [14, 15].

2.1 Complexidade Estatística de Permutação

Complexidade é um conceito difícil de se elaborar, mas a princípio, nossa intuição nos diz que algo "complexo" é algo que não é simples. Pode parecer óbvio mas, fora de nosso senso comum, poucas ferramentas são capazes de separar sistemas simples de complexos. Devido a própria dificuldade de definição de complexidade, várias medidas de complexidade foram sugeridas por pesquisadores, ilustrando o fato de que a dificuldade na definição de complexidade leva a várias tentativas diferentes de se definir o mesmo objeto [16].

A entropia de Bandt e Pompe [8] é uma dessas medidas de complexidade propostas, mas com as vantagens e desvantagens já discutidas anteriormente. Conforme discutido no artigo de Rosso *et al.* [9], a entropia de Bandt e Pompe apresenta a deficiência de ser incapaz de distinguir entre séries temporais caóticas e estocásticas. No entanto, o método introduzido pela entropia de permutação é poderoso e quando aliado à definição de complexidade proposta por López-Ruiz *et al.*. [10], temos um procedimento para a caracterização de séries temporais bastante refinado e completo, praticamente eliminando todas as deficiências da entropia de permutação.

Desse modo, López-Ruiz *et al.*. [10] propuseram uma definição de complexidade baseada em algumas noções intuitivas de física e estatística. Essas noções intuitivas são oriundas

de dois sistemas distintos: o gás ideal e o cristal perfeito. Ambos são modelos simples e com modelos teóricos sólidos, macroscopicamente previsíveis e, devido a isso, podem ser considerados de complexidade muito baixa.

O gás ideal é completamente desordenado e o sistema pode ser encontrado em qualquer um dos seus estados acessíveis com a mesma probabilidade. Uma grande quantidade de "informação" seria necessária para descrever o estado do sistema, pois todos os estados contribuem em partes iguais à informação armazenada no gás ideal. Assim, ele tem armazenado uma quantidade máxima de "informação".

Por outro lado, um cristal perfeito é completamente ordenado e os átomos são organizados seguindo regras de simetria. A distribuição de probabilidade para os estados acessíveis no cristal perfeito é centrado em um estado privilegiado de simetria. Uma quantia de "informação" pequena é suficiente para descrever o cristal perfeito: apenas as distâncias e as simetrias que definem a célula unitária são suficientes. Este sistema, portanto, armazena uma quantidade mínima de "informação". Esses dois exemplos são extremos na escala de ordem e "informação". Como a nossa definição de entropia está ligada à "informação", então o cristal perfeito tem entropia baixa enquanto o gás ideal tem entropia alta.

Assim, uma definição de complexidade deve ser capaz de englobar esses dois sistemas distintos. Analisando a distribuição de probabilidade desses dois sistemas, a primeira definição intuitiva seria sugerir uma medida de complexidade chamada "desequilíbrio" $\mathcal{D}$, no qual "desequilíbrio" é definido [10] como uma distância entre a distribuição equiprovável e a distribuição de probabilidade dos estados acessíveis ao sistema. Desse modo, a descrição de "desequilíbrio" dá a ideia de uma hierarquia entre os estados possíveis do sistema. Então, $\mathcal{D}$ será diferente de zero se existirem estados privilegiados ou mais prováveis dentre os estados acessíveis.

Porém, uma análise mais minuciosa mostra que definir $\mathcal{D}$ somente como essa distância não é suficiente. Por exemplo, no caso do do cristal perfeito temos um estado privilegiado que tem alta prioridade, segundo a ideia de hierarquia. A distância entre esse estado privilegiado e a distribuição equiprovável daria uma valor máximo para o desequilíbrio. Já o gás ideal tem desequilíbrio zero, visto que todos seus estados são equiprováveis. Assim o gás ideal tem complexidade $\mathcal{D} = 0$ por definição. Este resultado mostra que definir a complexidade apenas como a distância não é o suficiente. Ambos o gás ideal e o cristal perfeito deveriam apresentar complexidade baixa, mas esse último apresenta um resultado oposto ao esperado.

Porém, nós ainda podemos usar essa medida intuitiva de desequilíbrio. Analisando a figura 2.1, observamos o comportamento qualitativo esperado da entropia H (informação) e do desequilíbrio $\mathcal{D}$. Como mencionamos, nem H nem $\mathcal{D}$ são suficientes para descrever a complexidade, mas com base no comportamento dessas grandezas, López-Ruiz *et al.* [10] sugerem que a complexidade C seja definida como a multiplicação dessas duas quantidades, isto é, $C = H\mathcal{D}$. Essa função C tem o comportamento assintótico desejado, uma vez que ela

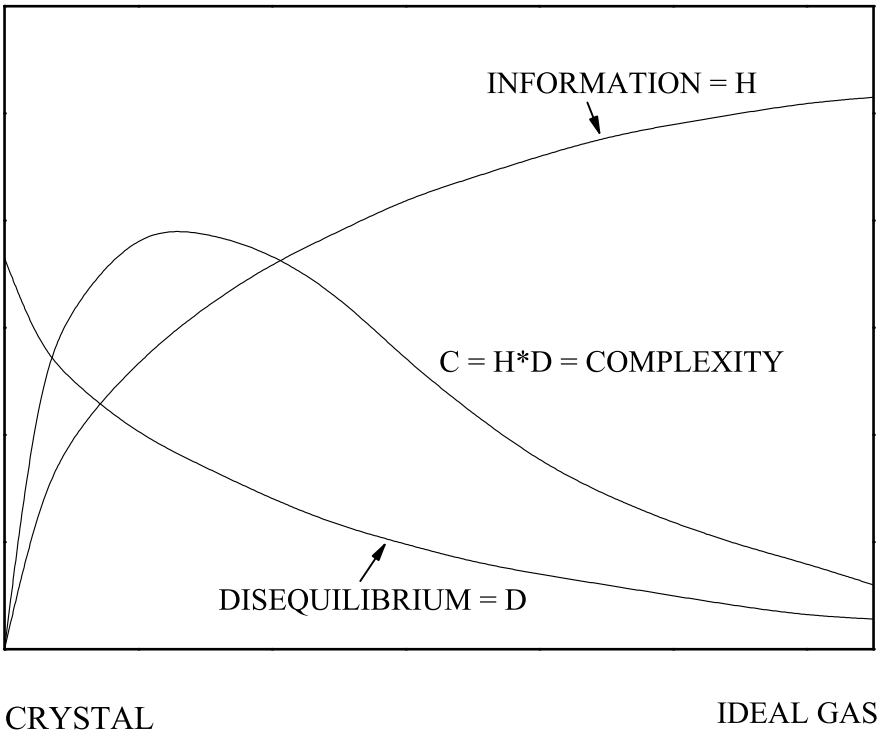

Figura 2.1: Representação do comportamento da entropia H (informação) e do desequilíbrio $\mathcal{D}$ para o gás ideal e o cristal perfeito. Sistemas ordenados possuem menos informação mas o desequilíbrio é máximo. A complexidade proposta por López-Ruiz *et al.* definida como $C = H \times \mathcal{D}$ é uma função que capta ambas as quantidades intuitivas requiridas, tendendo a zero no caso do gás ideal e do cristal perfeito. Figura adaptada da referência [17].

tende a zero para o gás ideal e para o cristal perfeito, mas é diferente de zero para outros sistemas que se encaixam nesses dois extremos.

Assim, a complexidade de López-Ruiz *et al.*. é definida matematicamente por

$$C[P] = H[P]\mathcal{D}[P] = -C_0 \left(\sum_{i=1}^{d!} p(\pi_i) \log p(\pi_i) \right) \left(\sum_{i=1}^{d!} \left(p(\pi_i) - \frac{1}{d!} \right)^2 \right), \qquad (2.1)$$

onde $H[P]$ é a entropia de Shannon, o desequilíbrio $\mathcal{D}[P]$ é a distância euclideana entre a distribuição de probabilidade e a distribuição equiprovável, C_0 é uma constante de normalização e $P = \{p(\pi_i)\}_{i=1,2,\ldots,d}$ é o conjunto de probabilidades dos estados acessíveis do sistema.

Retomando o exemplo do cristal e do gás ideal do começo dessa seção, vemos que essa medida de complexidade está de acordo com os resultados intuitivos esperados. Para um cristal perfeito o desequilíbrio $\mathcal{D}$ obtem um valor alto por causa da presença de um estado privilegiado, mas esse valor grande é multiplicado por um valor pequeno, a entropia H, que é pequena devido à baixa informação que o cristal carrega, de modo que $\mathcal{D}H \to 0$. No caso do gás ideal, o oposto acontece: a entropia H é grande pois a informação é máxima, o desequilíbrio $\mathcal{D}$ é mínimo devido à pequena distância entre a distribuição de probabilidade

do gás ideal e a distribuição equiprovável, e novamente a complexidade $\mathcal{D}H \to 0$.

Com a entropia e complexidade, podemos fazer um gráfico da entropia H pela complexidade $C(H)$. Entretanto, uma característica importante da complexidade proposta por López-Ruiz *et al.* [10] (equação 2.1) é que ela não é uma função unívoca da entropia. Isso significa que para um dado valor de entropia H vários valores diferentes da complexidade C são possíveis, variando entre o valor mínimo C_{min} e C_{max} ilustrado na figura 2.2. Visto de outro modo, podemos dizer que várias distribuições diferentes p_i armazenam a mesma informação H mas têm complexidade C diferentes. Usando um exemplo simples (adaptado da referência [18]) para analisar que entropias diferentes podem gerar a mesma complexidade. Consideramos um sistema que possui somente 3 estados acessíveis. A distribuição de probabilidade desses estados pode ser representada por $P = \{a, b, 1-(a+b)\}$ com $a > 0$, $b > 0$ e $(a+b) \leq 1$. A entropia para essas distribuições é dada por

$$H = -(a \log a + b \log b + [1-(a+b)] \log[1-(a+b)]),$$

e o desequilíbrio $\mathcal{D}$ fica

$$\mathcal{D} = \left(a - \frac{1}{3}\right)^2 + \left(b - \frac{1}{3}\right)^2 + \left([1-(a+b)] - \frac{1}{3}\right)^2.$$

Tomando os seguintes valores de $a, b, \text{e}\, c$

$$P_1 = \{0{,}79, 0{,}18, 0{,}03\} \text{ e } P_2 = \{0{,}80, 0{,}16, 0{,}04\},$$

podemos ver que ambos possuem os mesmos valores de H, evidenciando que a entropia H e C não têm uma correspondência um a um. Calculando a entropia, obtemos aproximadamente $H \approx 0{,}600$, mas valores diferentes de $\mathcal{D}$ e, consequentemente, de C. Para $\mathcal{D}$ temos 0,324 e 0,334, no primeiro caso, e C, 0,1944 e 0,2004, no segundo caso. Combinando todos os possíveis valores de a e b, podemos construir o plano complexidade-entropia observado na figura 2.2.

Até aqui utilizamos a entropia de Shannon na complexidade estatística e a distância euclideana para definir a complexidade. Mas como a entropia de Shannon pode ser calculada pelo método de Bandt e Pompe, com todas as suas vantagens, a aplicação da entropia de permutação na complexidade estatística de López-Ruiz *et al.* nos leva à "complexidade de permutação", que é o próximo refinamento do método utilizado para o estudo de séries temporais, como proposto por Rosso *et al.* [9].

Como detalhado anteriormente, utilizamos o método de Bandt e Pompe para encontrar a distribuição de probabilidade dos padrões ordinais $P = \{p(\pi_i)\}_{i=1,\ldots,d!}$ e então calculando

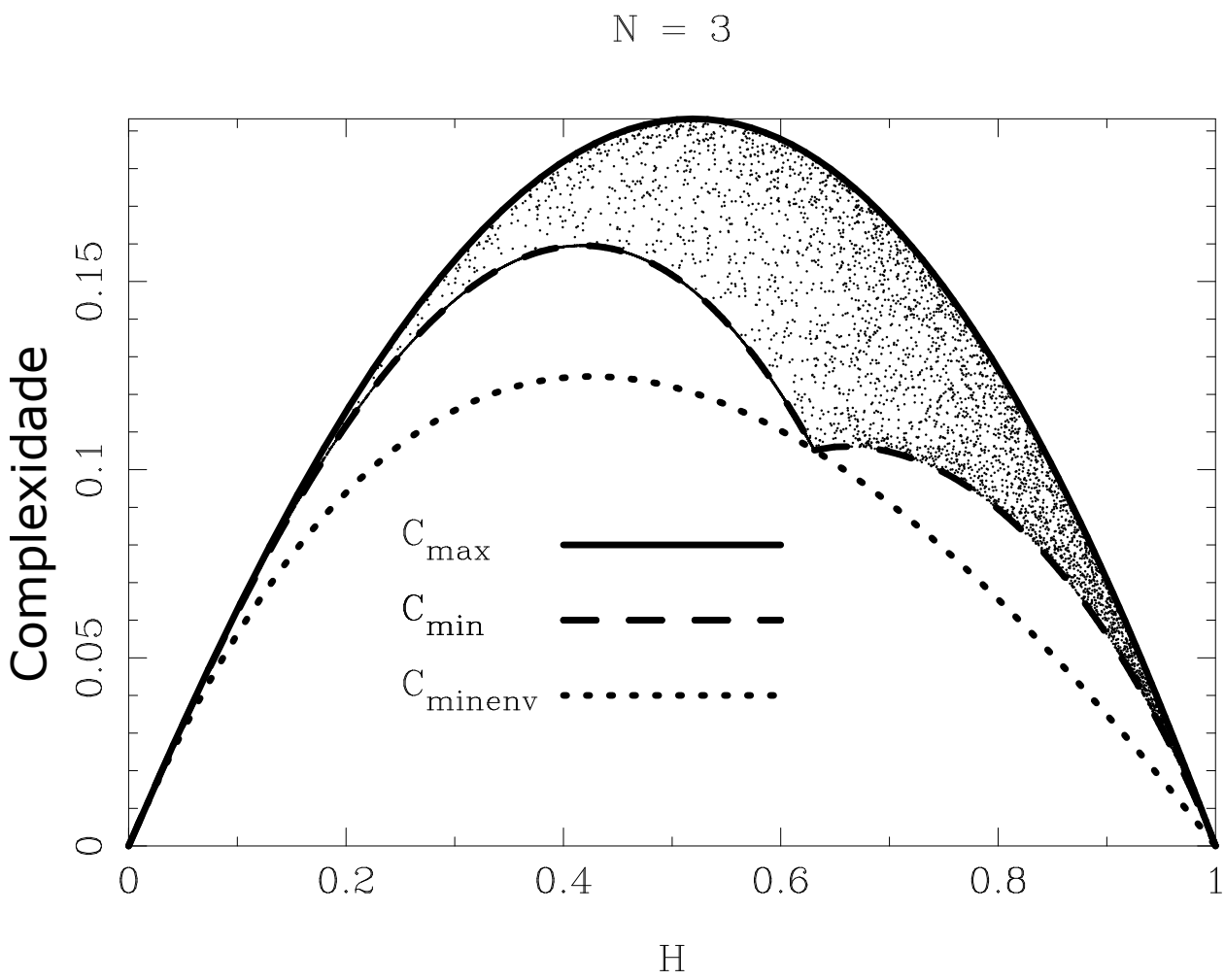

Figura 2.2: Plano complexidade-entropia para o exemplo com $N = 3$. Note que para cada valor de H temos vários valores de complexidade. Para o valor de $H \approx 0,7$ observamos que existem vários valores de C. O plano complexidade entropia é limitado superiormente por C_{max} e envelopado inferiormente por C_{minenv}. C_{min} é a curva que limita os pontos do nosso exemplo. Figura adaptada da referência [17]

a entropia de permutação normalizada,

$$H[P] = -\frac{1}{\log d!} \sum_{i=1}^{d!} p(\pi_i) \log p(\pi_i) \,, \tag{2.2}$$

e a complexidade estatística

$$C[P] = H[P]\mathcal{D}[P] \,. \tag{2.3}$$

Porém, em vez de usar a distância euclideana como no trabalho de López-Ruiz *et al.* [10], Rosso *et al.* [9] alteram o desequilíbrio $\mathcal{D}[P]$ para o desequilíbrio medido pela divergência de Jensen-Shannon [14, 15], ou seja,

$$\mathcal{D}[P] = \mathcal{D}_0 \left\{ H\left[\frac{P + P_e}{2}\right] - \frac{H[P]}{2} - \frac{H[P_e]}{2} \right\} . \tag{2.4}$$

Onde $P = \{p(\pi_i)\}_{i=1,\ldots,d!}$ é a distribuição de probabilidade, P_e é a distribuição de probabilidade equiprovável.

Observe que

$$\frac{P + P_e}{2} = \left\{ \frac{p(\pi_i) + 1/d!}{2} \right\}_{i=1,\ldots,d!} \tag{2.5}$$

E $\mathcal{D}_0$ é uma constante de normalização, obtido com o máximo valor de $\mathcal{D}[P]$, que acontece quando a probabilidade $P*$ que maximiza $\mathcal{D}$ é igual a um e todas as outras são zero tal que $P^* = \{\delta_{i1}\}_{i=1,\ldots,d!}$ [19]. Desse modo, a constante fica

$$\mathcal{D}_0 = \left\{ -\frac{1}{2} \left(\frac{d! + 1}{d!} \log(d! + 1) - 2 \log 2d! + \log d! \right) \right\}^{-1}. \tag{2.6}$$

Nesse contexto, o nome complexidade de permutação apenas reflete o uso conjunto de duas ferramentas: complexidade estatística com a entropia de permutação.

2.2 Outras medidas de complexidade

A complexidade definida por López-Ruiz *et al.* utiliza a entropia de Shannon e a distância euclidiana como sua entropia e distância, respectivamente. Mas a complexidade também pode ser obtida por meio de outras medidas entrópicas e medidas de desequilíbrio existentes [20, 21].

Entre as medidas de entropia, podemos listar:

- A já citada entropia de Shannon [4]

$$H[P] = -\sum_{i}^{d!} p(\pi_i) \log p(\pi_i); \tag{2.7}$$

- a entropia de Tsallis [22]:

$$H_q[P] = \frac{1}{q-1} \left[1 - \sum_{i}^{d!} (p(\pi_i))^q \right], \tag{2.8}$$

  onde q é um parâmetro real;

- e a entropia de Rényi [23]:

$$H_\alpha[P] = \frac{1}{1-\alpha} \log \left[\sum_{i}^{d!} (p(\pi_i))^\alpha \right], \quad (\alpha \geq 0 \text{ e } \alpha \neq 1), \tag{2.9}$$

  onde α é um parâmetro real;

Entre as medidas de distância/divergência, listamos:

- a distância euclidiana

$$\mathcal{D}[P] = \sum_{i}^{d!} \left(p(\pi_i) - \frac{1}{d!} \right)^2 ; \tag{2.10}$$

- a distância de Wooters [24]:

$$\mathcal{D}[P] = \cos^{-1} \left\{ \sum_{i}^{d!} (p(\pi_i))^{1/2} \left(\frac{1}{d!} \right)^{1/2} \right\}; \tag{2.11}$$

- e a divergência de Jensen-Shannon [14]:

$$\mathcal{D}[P] = \mathcal{D}_0 \left\{ H \left[\frac{P + P_e}{2} \right] - \frac{H[P]}{2} - \frac{H[P_e]}{2} \right\}. \tag{2.12}$$

  com H sendo a entropia de Shannon (equação 2.7) e $\mathcal{D}_0$ uma constante de normalização.

Embora não seja o foco do presente trabalho, vale ressaltar que essas diferentes definições para entropia, divergência e complexidade permitem explorar ou ressaltar aspectos diferentes da dinâmica de sistemas complexos [25, 26], além de contribuírem para exploração das diferentes ideias e definições sobre o conceito de complexidade.

2.3 Plano complexidade-entropia

A complexidade intuitiva introduzida por López-Ruiz *et al.* utiliza a entropia de Shannon como medida da quantidade de informação. Assim, Rosso *et al.* [9] sugere a utilização da entropia de permutação de Bandt e Pompe em conjunto com a complexidade estatística de López-Ruiz *et al.*, para calcular a chamada complexidade de permutação. Rosso *et al.* sugeriram em seu artigo [9] o uso da complexidade de permutação como um método para distinguir entre sinais caóticos e ruídos (sinais estocásticos), distinção que somente a entropia de permutação é incapaz realizar. Assim, eles introduzem a complexidade de permutação juntamente com o plano complexidade-entropia (ou *complexity-entropy plane*).

Esse diagrama, igual ao da figura 2.2, no qual a entropia é representada no eixo horizontal e a complexidade é representada no eixo vertical, é utilizado para distinguir as séries temporais caóticos das séries estocásticas, baseado na sua localização nesse diagrama. Esse plano também é de grande utilidade para analisar a dinâmica de sistemas complexos, o que pode ser feito, por exemplo, alterando algum parâmetro ou característica e seguindo os valores no plano $C \times H$.

Para aplicar o plano complexidade-entropia, Rosso *et al.* utilizaram várias séries temporais como teste para analisar e confirmar se são estocáticas ou caóticas. Na figura 2.3,

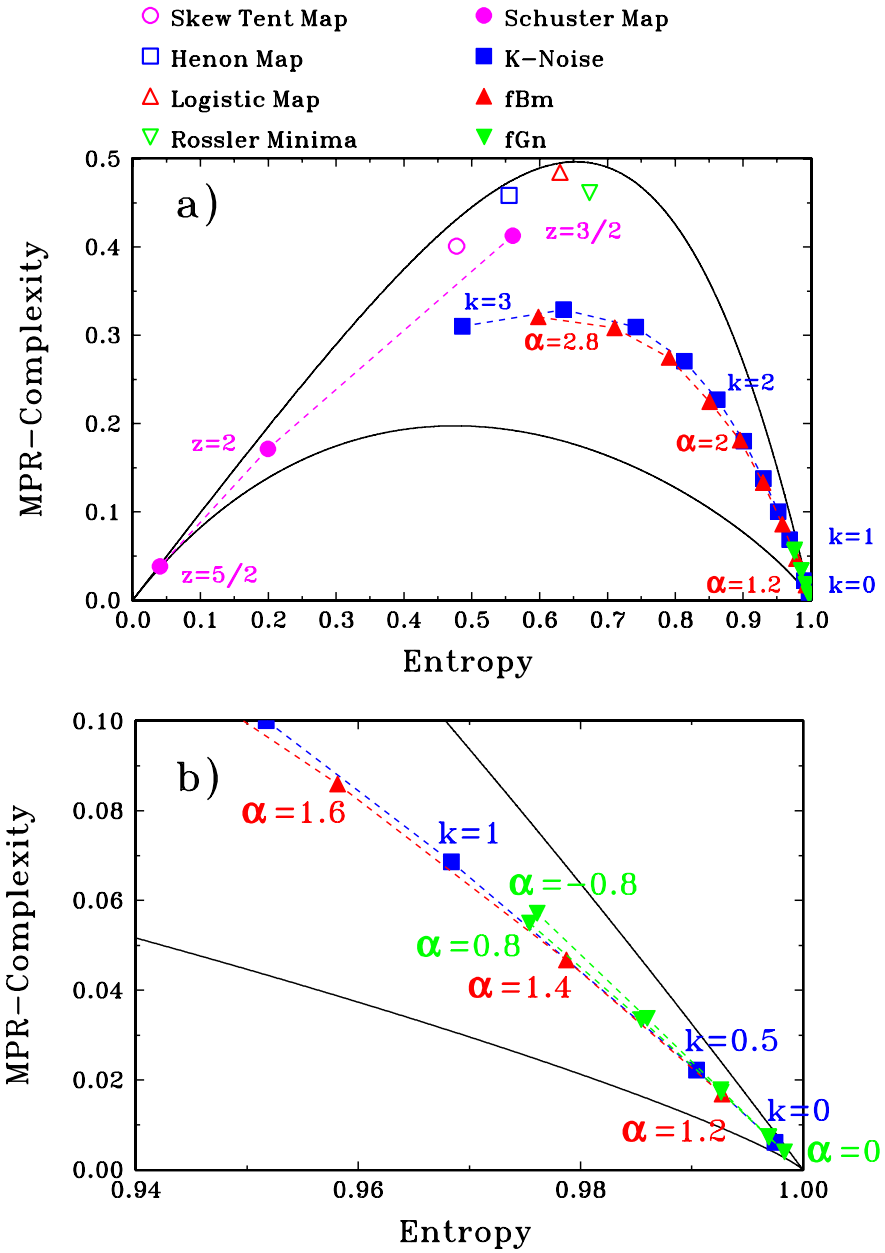

Figura 2.3: *Complexity-entropy plane.* O plano complexidade-entropia $C \times H$ é delimitado pelas linhas contínuas, que representam a complexidade máxima C_{max} e mínima C_{min}. Os símbolos nesse gráfico indicam a localização de diferentes sinais caóticos (os cinco primeiros da legenda) e estocásticos (os três últimos) nesse plano. Em (a) Percebemos que esses sinais são facilmente distinguíveis, mesmo quando possuem a mesma entropia. Por exemplo, séries estocásticas do ruído-k podem ter a mesma entropia de séries caóticas do mapa logístico, porém complexidades completamente diferentes. (b) Zoom na região de máxima entropia $H \approx 1$, mostrando como as séries estocásticas evoluem e se aproximam da região de completa aleatoriedade. Figura adaptada da referência [9].

estão os resultados obtidos e seus valores representados no plano. As cinco primeiras séries: tenda inclinada, Henon, Logístico, mínimos de Rossler e Schuster são todas séries caóticas oriundas de mapas (representados pelos ícones vazios e o círculo cheio). Os três últimos:

ruído-k, movimento browniano fracionário e ruído gaussiano fracionário, são séries estocásticas (aleatórias). Podemos notar que séries temporais caóticas e estocásticas ocupam regiões diferentes do plano complexidade-entropia, tornando fácil a distinção entre esses diferentes sistemas.

De maneira geral, Rosso *et al.* [9] notaram que sistemas caóticos estão localizados próximos da região de máxima complexidade e com entropia na região entre 0.45 e 0.7. O motivo para isso, os autores argumentam, é porque esses sistemas possuem um grau maior de estruturas "imersas" em suas dinâmicas, como a ausência de certos padrões de permutação. Assim, notamos que a complexidade estatística quantifica, além da aleatoriedade, as dinâmicas naturais existentes que geram a série temporal. Vale notar ainda que, sozinhas, a entropia de permutação ou a complexidade estatística não são capazes de fazer essa distinção por completo. Por exemplo, séries temporais do movimento browniano fracionário podem apresentar a mesma entropia de séries do mapa logistíco, mas complexidades diferentes. Similarmente, o mapa de Schuster para $z = 2$ tem complexidade similar a do movimento browniano fracionário, mas o valor da entropia é muito diferente.

2.4 Aplicações do plano complexidade-entropia à series empíricas

Uma outra aplicação desse diagrama que estudamos está relacionada a séries temporais da amplitude sonora de músicas de vários gêneros musicais [19, 27]. Nesse estudo, os autores usam a entropia e complexidade de permutação para tentar organizar os gêneros de acordo com sua "complexidade". Esse tipo de abordagem leva a uma classificação de músicas e seus gêneros dependendo dos padrões ordinais observados para as amplitudes sonoras.

O procedimento usado envolveu analisar séries temporais de amplitudes sonoras e intensidades sonoras de mais de 10 mil músicas e então representar cada uma delas no plano complexidade entropia. A figura 2.4 mostra a média dos valores de complexidade e entropia das músicas agrupadas por gênero musical. Note que ambas as séries de amplitudes e intensidades apresentam resultados similares.

O resultado também mostra a formação de grupos distintos de gêneros musicais, no qual gêneros de alto padrão estético como jazz e classical apresentam menor entropia e maior complexidade, representando as complexidades intrínsecas desses genêros que aparentemente não estão presentes nos demais estilos. Gêneros musicais como techno e pop estão mais próximos da região de permutações aleatórias ($H = 1$ e $C = 0$), indicando que o padrão ordinal dessas músicas são mais "pobres", no sentido de serem mais próximas do aleatório.

Uma outra análise realizada foi a quantificação da evolução da complexidade das músicas com o passar do tempo [27]. A motivação para essa análise foi a percepção de muitas

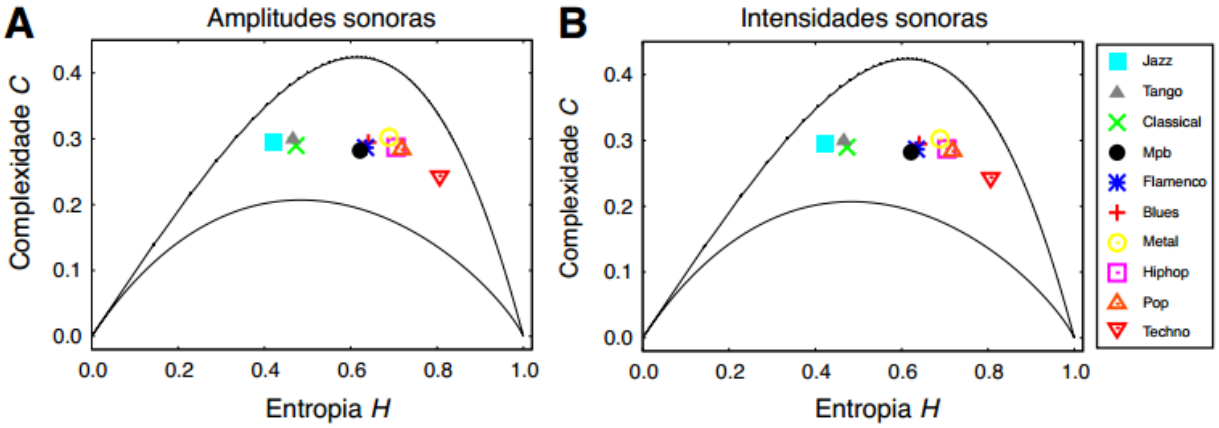

Figura 2.4: Média dos valores de entropia e complexidade para cada gênero musical usuando as (**A**) amplitudes sonoras e as (**B**) intensidades sonoras. Notamos ainda a existência de 4 grupos relativamente bem definidos: *G1*: jazz, tango e classical; *G2*: mpb, flamenco e blues; *G3*: metal, hiphop e pop; *G4*: techno. Figura adaptada da referência [27].

pessoas de que a maioria das músicas populares tiveram um declínio em "qualidade". A metodologia para a investigação desse possível aumento da "pobreza" nas música populares foi a seguinte: criar uma base de dados com as músicas mais populares nos Estados Unidos seguindo classificação da revista "Billboard", no período de 1946 a 2007. Anualmente, essa revista publica uma lista com as "top 50" músicas mais populares de cada ano. Usando essa base formada pelas músicas mais populares de cada ano, composta por mais de 5 mil músicas, foram calculadas a entropia e complexidade de permutação para as amplitudes sonoras de cada música, assim como seus valores médios em função dos anos, os quais estão representados na figura 2.5.

Nessa figura, percebemos um aumento da entropia e uma diminuição da complexidade com o passar dos anos. Além disso, a evolução desses índices parece se aproximar do limiar de permutações aleatórias ($H = 1$ e $C = 0$). Isso significa que os padrões ordinais contidos nas músicas tem se tornado menos complexos e mais aleatórios com o passar dos anos. Apesar de ser uma abordagem simplificada e reducionista, essa análise aponta que as músicas estão em constante evolução do ponto qualitativo. Apesar de não haver uma análise sobre os gêneros musicas a que estas músicas pertencem, uma das prováveis causas para esse decrescimento em complexidade se deve à popularização de gêneros e músicas mais dançantes, como pop e techno.

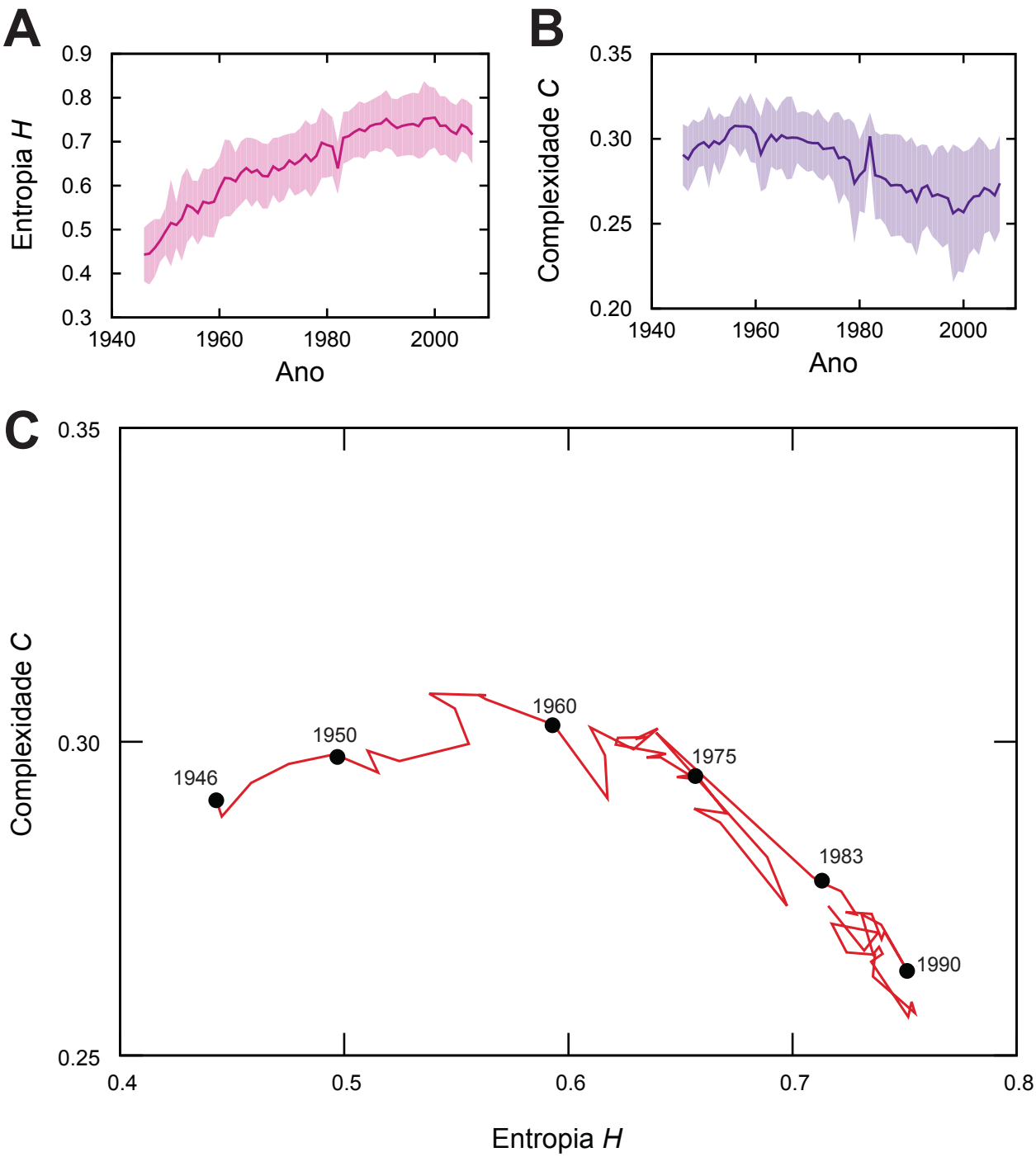

Figura 2.5: Alteração do padrão das músicas mais populares nos Estados Unidos (de acordo com a *Billboard and Top Charts*). (**A**) Evolução da entropia H e da (**B**) complexidade estatística de permutação ano a ano. É possível notar um claro aumento da entropia e diminuição da complexidade. Em (**C**) temos a evolução conjunta dos valores de C e H, na qual observamos que com o passar dos anos a entropia e complexidade tem se aproximado do limiar de permutações aleatórias. Figura adaptada da referência [27].

Outra aplicação da análise do plano complexidade-entropia que estudamos foi a análise da chamada ineficiência de mercado (*stock market inneficiency*) realizada por Zunino *et al.* [28]. A ineficiência de mercado é uma consequência da hipótese de mercado eficiente, que foi introduzida por Eugene Fama em 1970 [29], quando ele sugeriu que mercados financeiros são eficientes com relação à informação, ou seja, um agente não consegue alcançar retornos superiores a média do mercado consistentemente, considerando as informações publicamente disponíveis no momento que o investimento é feito.

Uma consequência disso é que seria impossível investidores comprarem ações subvalorizadas ou vender ações a preços inflados. Assim, seria impossível superar o mercado por meio de seleção esperta de ações ou de *timing* e o único jeito de alcançar lucro seria por puro acaso ou pela compra investimentos que valorizasem por si próprios.

A *Efficient Market Hypothesis* [30] é usualmente separada em três variações, cada uma delas com características relacionadas à informação que o preço do mercado de ações carrega e como elas afetam preços futuros. São elas:

- hipótese fraca: os preços negociados para os bens (por exemplo ações ou propriedades) refletem toda a informação histórica disponível publicamente. Preços futuros não podem ser previstos analisando preços do passado;
- hipótese semi-forte: os preços refletem as informações publicamente disponíveis e também os preços mudam instantaneamente para refletir as novas informações públicas;
- hipótese forte: os preços refletem as informações públicas instantaneamente e os preços também refletem informações ocultas ou privilegiadas (informações não-públicas).

A versão fraca da hipotese é apenas uma primeira aproximação. A existência de autocorrelação entre observações distantes quebra a hipótese fraca pois, desse modo, preços passados podem ajudar a prever preços futuros. Beben e Orłowski [31] e Di Matteo *et al.* [32,33] descobriram que mercados de ações de mercados emergentes têm maior correlação do que mercados desenvolvidos. Assim, mercados emergentes são menos eficientes do que mercados desenvolvidos.

Hoje em dia costuma-se assumir que uma ineficiência residual está sempre presente em mercados de ações e o conceito de eficiência de mercado é só uma idealização [28]. Conceitos como estabilidade política, perfil de risco e gerenciamento econômico distinguem mercados emergentes de mercados desenvolvidos. Além disso, liquidez e capitalização de mercado influenciam muito na eficiência do mercado; porém, essas quantidades e correlações são difíceis de se quantizar de maneira precisa.

Assim, mercados eficientes tendem a incorporar mais as informações disponíveis, tornando seus valores menos previsíveis devido à autocorrelação. No oposto disso, mercados ineficientes falham em incorporar mais informações, tornando seus valores mais previsíveis devido à autocorrelação.

Zunino *et al.* [28] propõem utilizar o plano complexidade-entropia em séries temporais do índice de bolsas de valores pois, como discutido anteriormente, juntas, a entropia de permutação e a complexidade estatística conseguem extrair e quantificar as estruturas "imersas" em suas dinâmicas, neste caso, as várias estruturas que correlacionam as informações e os preços das bolsas de valores.

Nesse trabalho, os dados utilizados foram o valor diário do índice das bolsas de valores de 32 países diferentes, no período de 1995 a 2007, acumulando um total de 3138 observações para cada bolsa de valores. Desses 32 países, 18 mercados de ações são considerados desenvolvidos, e 14 são considerados emergentes. Os valores para *embedding dimension* foram $d = 4, 5, 6$, pois o número de observações em cada série temporal limita o valor de d para que a condição seja satisfeita ($3138 >> 6! = 720$). O plano complexidade-entropia pode ser visto na figura 2.6.

De imediato, percebemos o plano complexidade-entropia é capaz de distinguir entre os mercados de países desenvolvidos e em desenvolvimento, ainda que alguma superposição exista. Conforme argumenta Zunino *et al.*, esses casos que se superpõem compreendem uma mistura de países desenvolvidos e emergentes (Argentina, Áustria, Canadá, Brasil, China, Grécia, Hong Kong, Singapura, Taiwan e Turquia). Tais países, apesar de serem classificados (de acordo com a metodologia para definir mercados emergentes e desenvolvidos do Morgan Stanley Capital Index, MSCI) como emergentes ou desenvolvidos, estão em uma região próxima entre si, mas claramente distinta dos outros países. Zunino *et al.* concluem que uma terceira categoria de países poderia ser sugerida: a de países *híbridos*. Essa categoria de países híbridos possuem valores de entropia e complexidade intermediários e compreendem 5 países desenvolvidos e 5 emergentes. Um padrão notado é que os países emergentes desse bloco híbrido estão em desenvolvimento e apresentam crescimento rápido e seu mercado, se tornando mais eficiente com o tempo (Argentina, China, Brasil, Taiwan e Turquia). Já os países desenvolvidos dessa lista, são relativamente pequenos se comparados outras economias desenvolvidas, como Estados Unidos, Reino Unido e Japão.

Também é possível concluir que mercados emergentes têm menor entropia e maior complexidade e algum grau de ordem, revelando a presença de autocorrelações de maior significância do que mercados desenvolvidos. Portanto, esses países apresentam mercados menos eficientes do ponto de vista informacional. Por outro lado, mercados desenvolvidos possuem maior entropia e menor complexidade, aproximando-se de um processo aleatório e revelando a presença de menos correlações e, portanto, uma maior eficiência informacional.

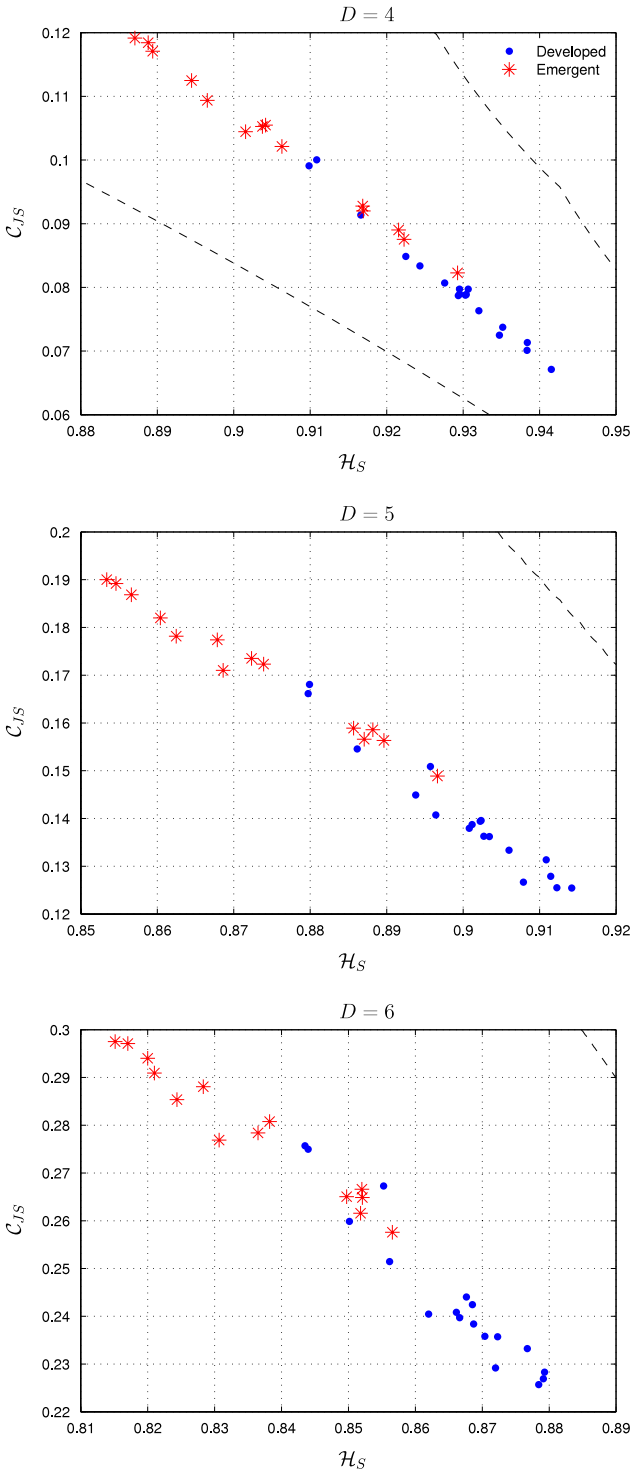

Figura 2.6: Localização dos mercados desenvolvidos e emergentes no plano complexidade-entropia, com $d = 4$ (topo), $d = 5$ (meio) e $d = 6$ (baixo). Também é mostrado os valores máximos e mínimos da complexidade estatística (linha tracejada). Em $d = 6$ notamos a presença de 3 grupos distintos: mercados emergentes, mercados desenvolvidos, e o caso intermediário composto de países emergentes e desenvolvidos (na região de H variando entre 0,84 e 0,86 e C entre 0,25 e 0,28). Figura adaptada da referência [28].

Conclusões e perspectivas

Nesse trabalho, revisamos as técnicas de entropia de permutação de Bandt e Pompe [8], a complexidade estatística de López-Ruiz et al. [10], e por fim, a complexidade de permutação o plano complexidade entropia desenvolvido por Rosso et al. [9].

Motivados pela dificuldade na caracterização de séries temporais, e na tentativa de distinguir séries caóticas de estocásticas, cada um dos métodos mostrou-se essencial no desenvolvimento e no refinamento do processo, no final conseguimos analisar com sucesso séries variadas e distingui-las de acordo com sua complexidade, além de conseguir definir com sucesso um parâmetro para distinguir séries estocásticas de séries caóticas. Por fim, estudamos alguns exemplos do uso dessas técnicas para a investigação de sistemas complexos reais.

Como perspectivas de estudos futuros, podemos listar o uso dessas técnicas baseadas em outras medidas de entropia e outras divergências estatísticas. Outra possibilidade de futuros estudos incluem as generalizações da técnica para sistemas multidimensionais, a qual permite a investigação de imagens e figuras, além das generalizações associadas ao uso múltiplas escalas temporais. Sendo assim, acreditamos que a revisão apresentada aqui pode servir de base para o novos estudos, sejam eles relacionados ao uso dessas ferramentas para a investigação de novos sistemas ou mesmo associados ao aprimoramento e extensão das técnicas de entropia e complexida de permutação.

Apêndice A: Cálculo de S_{max} via multiplicadores de Lagrange

Cálculos adaptados da referência [18].

A distribuição que maximiza a entropia de permutação tem que ser não negativa para quaisquer estado π_i, ou seja,

$$p(\pi_i) \geq 0 \, (\forall \, i), \tag{2.13}$$

além de ser normalizada, isto é,

$$\sum_{i=1}^{d!} p(\pi_i) = 1, \tag{2.14}$$

pois a probabilidade de encontrarmos o sistema pelo menos em um dos $d!$ possíveis estados é diferente de zero, e a soma de todas as probabilidades corresponde a 100%, ou seja, sempre encontraremos pelo menos um estado. Esses sãos os nossos vínculos.

Utilizando a equação dos multiplicadores de Lagrange, a função a ser maximizada com o vínculo é dada por

$$\mathcal{L}(p(\pi_i), \lambda) = -\sum_{i=1}^{d!} p(\pi_i) \log p(\pi_i) + \lambda \left[\sum_{i=1}^{d!} p(\pi_i) - 1 \right], \tag{2.15}$$

em que λ é o multiplicador de Lagrange.

Derivando essa equação para encontrar os valores extremos e igualando a zero, temos

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial p(\pi_j)} = 0 \implies & -\log p(\pi_j) - 1 + \lambda = 0, \\ & \log p(\pi_j) = \lambda - 1, \\ & p(\pi_j) = e^{\lambda - 1}, \end{aligned} \tag{2.16}$$

ou seja, as probabilidades $p(\pi_j)$ não dependem de j.

Realizando a soma para todos os estados, temos

$$\sum_{j=1}^{d!} p(\pi_j) = \sum_{j=1}^{d!} e^{\lambda-1}\,, \tag{2.17}$$

na qual o lado esquero é igual a 1 devido ao vínculo definido pela equação 2.14 e $e^{\lambda-1}$ independe de j, ou seja,

$$\begin{aligned} 1 &= e^{\lambda-1} \sum_{j=1}^{d!} 1, \\ \frac{1}{d!} &= e^{\lambda-1}. \end{aligned} \tag{2.18}$$

Comparando essa última equação com a última equação 2.16, concluímos que

$$p(\pi_j) = \frac{1}{d!}, \tag{2.19}$$

Desse modo, a distribuição de probabilidade "menos informativa" e que maximiza a entropia é

$$P_e = \{p(\pi_i) = 1/d!\,,\ i = 1, \ldots, d!\}, \tag{2.20}$$

a qual é a distribuição uniforme.

Calculando a entropia de Shannon para essa distribuição de probabilidade, obtemos

$$\begin{aligned} S_{max} = S[P_e] &= -\sum_{i=1}^{d!} \left(\frac{1}{d!}\right) \log\left(\frac{1}{d!}\right) \\ &= -\left(\frac{1}{d!}\right) \log\left(\frac{1}{d!}\right) \sum_{i=1}^{d!} 1 \\ &= -\left(\frac{1}{d!}\right) (\cancelto{0}{\log 1} - \log d!)\, d! \\ &= \log d!. \end{aligned} \tag{2.21}$$

Referências Bibliográficas

[1] Mantegna, R. N. & Stanley, H. E. *An Introduction to Econophysics: Correlations and Complexity in Finance*, vol. 53 (Cambridge University Press, 2000).

[2] Boccara, N. *Modeling complex systems* (Springer-Verlag, New York, 2004).

[3] Kolmogorov, A. N. Three approaches to the quantitative definition of information. *International Journal of Computer Mathematics* **2**, 157–168 (1968).

[4] Shannon, C. E. A mathematical theory of communication. *Bell System Technical Journal* **27**, 379–423 (1948).

[5] Kullback, S. & Leibler, R. A. On information and sufficiency. *The Annals of Mathematical Statistics* **22**, 79–86 (1951).

[6] Mandelbrot, B. B. *The Fractal Geometry of Nature* (W. H. Freeman, San Francisco, 1982).

[7] Lyapunov, A. M. The general problem of the stability of motion. *International Journal of Control* **55**, 531–534 (1992).

[8] Bandt, C. & Pompe, B. Permutation entropy: A natural complexity measure for time series. *Physical Review Letters* **88**, 174102 (2002).

[9] Rosso, O. A., Larrondo, H. A., Martin, M. T., Plastino, A. & Fuentes, M. A. Distinguishing noise from chaos. *Physical Review Letters* **99**, 154102 (2007).

[10] López-Ruiz, R., Mancini, H. L. & Calbet, X. A statistical measure of complexity. *Physics Letters A* **209**, 321–326 (1995).

[11] Zanin, M., Zunino, L., Rosso, O. & Papo, D. Permutation entropy and its main biomedical and econophysics applications: A review **14**, 1553 (2012).

[12] Courant, R. *Differential and Integral Calculus, Vol. 2* (Wiley-Interscience, New York, 1988).

[13] Rabiner, L. R. & Schafer, R. W. *Digital Processing of Speech Signals* (Prentice-Hall, 1978).

[14] Grosse, I. *et al.* Analysis of symbolic sequences using the Jensen-Shannon divergence. *Physical Review E* **65**, 041905 (2002).

[15] Lin, J. Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory* **37**, 145–151 (1991).

[16] Ribeiro, H. V., Jauregui, M., Zunino, L. & Lenzi, E. K. Characterizing time series via complexity-entropy curves. *Phys. Rev. E* **95**, 062106 (2017).

[17] Ricardo Lopez-Ruiz, X. C., Hector Mancini. *Concepts and Recent Advances in Generalized Information Measures and Statistics* (Bentham Science).

[18] Sigaki, H. Y. D. Física estatística aplicada ao estudo de obras de arte. *Universidade Estadual de Maringá - Tese de Mestrado* (2017).

[19] Ribeiro, H. V., Zunino, L., Mendes, R. S. & Lenzi, E. K. Complexity–entropy causality plane: A useful approach for distinguishing songs. *Physica A: Statistical Mechanics and its Applications* **391**, 2421 – 2428 (2012).

[20] Martin, M. T., Plastino, A. & Rosso, O. A. Statistical complexity and disequilibrium. *Physics Letters A* **311**, 126–132 (2003).

[21] Kowalski, A. M., Martín, M. T., Plastino, A., Rosso, O. A. & Casas, M. Distances in probability space and the statistical complexity setup. *Entropy* **13**, 1055–1075 (2011).

[22] Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics* **52**, 479–487 (1988).

[23] Rényi, A. On measures of entropy and information (The Regents of the University of California, 1961).

[24] Wootters, W. K. Statistical distance and Hilbert space. *Physical Review D* **23**, 357–362 (1981).

[25] Ribeiro, H. V., Zunino, L., Lenzi, E. K., Santoro, P. A. & Mendes, R. S. Complexity-entropy causality plane as a complexity measure for two-dimensional patterns. *PLOS ONE* **7**, e40689 (2012).

[26] Jauregui, M., Zunino, L., Lenzi, E., Mendes, R. & Ribeiro, H. Characterization of time series via rényi complexity–entropy curves. *Physica A: Statistical Mechanics and its Applications* **498**, 74 – 85 (2018).

[27] Ribeiro, H. V. Identificação e modelagem de padrões em sistemas complexos. *Universidade Estadual de Maringá - Tese de Doutorado* (2012).

[28] Zunino, L., Zanin, M., Tabak, B. M., Pérez, D. G. & Rosso, O. A. Complexity-entropy causality plane: A useful approach to quantify the stock market inefficiency. *Physica A: Statistical Mechanics and its Applications* **389**, 1891 – 1901 (2010).

[29] Fama, E. F. Efficient capital markets: A review of theory and empirical work. *Journal of Finance* **25**, 383–417 (1970).

[30] Malkiel, B. *A Random Walk down Wall Street* (W. W. Norton Company, Inc,, 1973).

[31] Beben, M. & Orłowski, A. Correlations in financial time series: established versus emerging markets. *The European Physical Journal B - Condensed Matter and Complex Systems* **20**, 527–530 (2001).

[32] Di Matteo, T., Aste, T. & Dacorogna, M. Scaling behaviors in differently developed markets. *Physica A: Statistical Mechanics and its Applications* **324**, 183–188 (2003).

[33] Di Matteo, T., Aste, T. & Dacorogna, M. Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development. *Journal of Banking and Finance* **29**, 827–851 (2004).