Problemas Interessantes

July 4, 2021

1 Estudo de padrões ordinais com grafos de transição e Cadeias de Markov

Uma das limitações das técnicas de análise de séries temporais baseadas em padrões ordinais é não permitirem fazer previsões (forecast) nem simulações. Este trabalho explora essa possibilidade.

Sejam uma série temporal de valores reais, a série de padrões a ela associados (calculados com palavras de dimensão e atraso), e o grafo de transições obtido a partir de .

A proposta consiste em fazer a junção das evidências de , , e para fazer simulações e previsões a respeito da série.

O primeiro passo consiste em analisar e . Para cada padrão observado em , serão coletados os dados que o originaram em . Consideremos, por exemplo, o caso do padrão , e suponhamos que ele corresponde a todas as palavras que satisfazem . Todas essas palavras serão coletadas e analisadas para obter:

- uma estimativa da distribuição três-variada, ou
- estimativas do valor central e estimativas de uma medida de dispersão de cada um dos três valores, por exemplo a média e o desvio padrão.

Teremos, assim, associados ao padrão ,

- um modelo , ou
- três médias e três desvios padrão .

O segundo passo consiste em formar a matriz de transições do grafo , digamos . Por construção, a cadeia é irredutível, e basta com que haja uma única transição entre estados iguais para que a cadeia seja aperiódica. Com estas propriedades, há uma única distribuição de equilíbrio , que é a solução de .

Dada a série temporal , associada à sequência de padrões, simularemos o evento com dois elementos:

- o padrão que possui probabilidade máxima de ocorrência em após o último padrão, e
- uma observação do modelo correspondente a esse padrão. Note-se que será necessário obter apenas uma amostra da distribuição marginal de dadas as observações já presentes nos últimos estágios da série.

A nossa previsão da observação que sucede será o estado de equilíbrio mais plausível que segue ao último padrão que inclui , e em cada posição do padrão colocaremos a estimativa de centralidade, munida da sua estimativa de precisão. De forma mais sofisticada, poderemos usar o algoritmo EM (?).

2 Desvios do Ponto

Seja uma sequência de observações de variáveis aleatórias independentes e identicamente distribuídas segundo uma lei , e a sequência de padrões ordinais de e dimensão . Seja ainda o histograma de proporções dos padrões .

A entropia de Shannon do histograma de proporções

Utilizaremos como distribuição de equilíbrio a lei uniforme sobre o conjunto , que é caracterizada pelas probabilidades . Denotaremos . A entropia de é .

A distância de Jensen-Shannon entre e é

Suponhamos que , , com , e que para todo . Com essa hipótese

Verificamos que

(4)

(5)

Assim, temos

Com (5) e (??) calculamos a complexidade estatística

A Fig. 1 mostra os desvios do ponto , para , e variações de entre e .

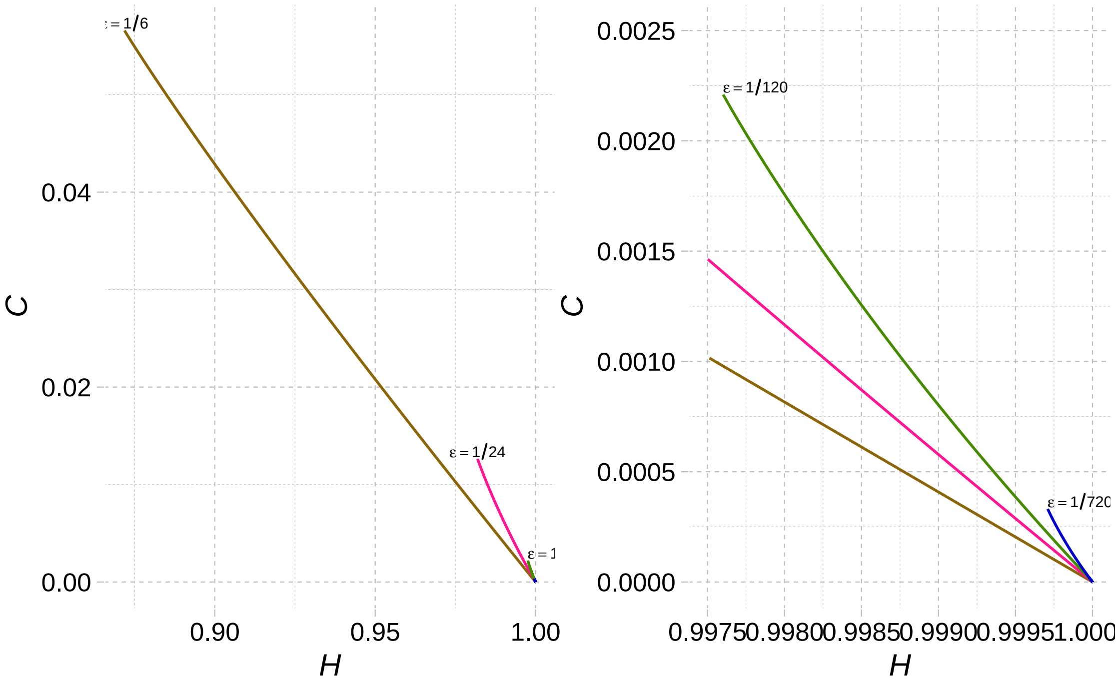

Figure 1: Desvios do ponto ao considerar que um par de bins foi alterado, para .