

## Diário de Pesquisa

### Dia 1: Exploração Inicial da Teoria da Informação e Entropia de Permutação

Data: 5 de Outubro de 2024

Hoje iniciei meus estudos aprofundados em **teoria da informação** e entropias aplicadas a séries temporais. O foco foi entender os conceitos fundamentais de **entropia de permutação** e como ela se relaciona com a **informação em sistemas dinâmicos**. Também explorei alguns capítulos iniciais de livros como "**Elements of Information Theory**" (Cover & Thomas) e "**Information Theory and Reliable Communication**" (Gallager).

#### Aprendizados do Dia 1:

- Entropia de Shannon vs. Entropia de Permutação:** A entropia de Shannon mede a incerteza média associada à ocorrência de eventos em um sistema probabilístico. A **entropia de permutação**, por outro lado, é focada na ordem relativa dos valores em séries temporais, capturando a complexidade baseada na dinâmica subjacente dos dados.
- Séries Temporais e Estrutura:** Em uma série temporal com forte estrutura (como uma tendência clara), a entropia de permutação é mais baixa, pois a ordem dos valores é mais previsível. Já em séries caóticas ou ruidosas, a entropia de permutação tende a ser mais alta, indicando maior complexidade.
- Primeiros Passos no Cálculo:** A entropia de permutação é calculada organizando subsequências de uma série temporal e analisando a ordem relativa dos valores em cada subsequência.

#### Exercícios Resolvidos:

##### Exercício (avulso):

Calcule a entropia de um dado justo de seis lados.

##### Solução:

A entropia de Shannon é dada por:

$$H(X) = - \sum_{i=1}^n p_i \log_2 p_i$$
 Para um dado justo,  $p_i = \frac{1}{6}$  para todos os lados. Assim: 
$$H(X) = -6 \times \left( \frac{1}{6} \log_2 \frac{1}{6} \right) = \log_2 6 \approx 2.585 \text{ bits}$$

##### Exercício (avulso):

Explique a diferença entre entropia e informação mútua.

##### Solução:

A **entropia** mede a incerteza total em uma variável aleatória. Já a **informação mútua** mede a dependência entre duas variáveis, ou seja, o quanto saber o valor de uma variável reduz a incerteza sobre a outra.

### Dia 2: Experimentação

Data: 7 de Outubro de 2024

Hoje avancei para a parte prática, onde comecei a **implementar o cálculo de entropia de permutação** em séries temporais sintéticas e reais. Meu objetivo foi entender como a escolha dos parâmetros (dimensão de embedding e atraso) afeta a precisão e o significado dos resultados. Também fiquei curioso com aplicação em caso aplicados no mercados financeiros.

## Aprendizados do Dia 2:

### 1. Escolha dos Parâmetros:

2. **Dimensão de Embedding (m):** Determina o número de pontos usados para formar subsequências na série. Valores muito baixos podem não capturar toda a complexidade, enquanto valores muito altos tornam o cálculo sensível ao ruído.

3. **Lag ( $\tau$ ):** O espaçamento entre os pontos nas subsequências. Valores muito altos podem perder a correlação entre pontos próximos, enquanto valores muito baixos podem não capturar dinâmicas de longo prazo.

### 4. Implementação do Algoritmo:

Utilizei **Python** para implementar o cálculo. Aqui está um trecho de código simplificado: ``python  
import numpy as np from pyentrp import entropy as ent

```
# Exemplo de série temporal sintética serie = np.random.rand(100)
```

```
# Cálculo da entropia de permutação entropia_perm = ent.permutation_entropy(serie, m=3, delay=1,  
normalize=True) print(f'Entropia de Permutação: {entropia_perm}') ``
```

---

## Exercícios Resolvidos:

### Exercício (avulso):

*Explique a importância da redundância em um sistema de comunicação.*

### Solução:

A **redundância** é essencial para garantir a **confiabilidade** em sistemas de comunicação. Ela permite que erros sejam detectados e corrigidos, aumentando a integridade da mensagem transmitida, mesmo em ambientes ruidosos.

---

## Dia 3: Introdução à Inferência Estatística Baseada em Medidas de Divergência

**Data:** 10 de Outubro de 2024

Hoje comecei a leitura dos primeiros capítulos do livro "**Statistical Inference based on Divergence Measures**", de Leandro Pardo. A introdução foca nas motivações para o uso de medidas de divergência na inferência estatística, destacando como essas medidas generalizam conceitos como a divergência de Kullback-Leibler e oferecem maior flexibilidade em modelos não tradicionais.

---

## Aprendizados do Dia 3:

### 1. Medidas de Divergência:

Essas medidas generalizam a noção de "distância" entre uma distribuição verdadeira e uma distribuição teórica. A divergência de Kullback-Leibler, amplamente utilizada, é apenas um caso particular.

### 2. Aplicação na Inferência:

Uma inferência baseada em divergência minimiza essa distância entre a distribuição observada e a esperada, tornando-a especialmente útil em situações onde suposições clássicas, como normalidade, não se mantêm.

---

## Dia 4: Exploração da Inferência de Mínima Distância

**Data:** 12 de Outubro de 2024

Hoje estudei o capítulo introdutório do livro "**Statistical Inference: The Minimum Distance Approach**" de Basu, Shioya e Park. A inferência por distância mínima é apresentada como uma abordagem robusta e flexível para diversos tipos de modelos.

---

#### Aprendizados do Dia 4:

1. **Distância Quadrática e Outras Funções:**

A inferência baseada em distância mínima minimiza uma função de distância entre a distribuição empírica e a teórica. Isso é útil em problemas robustos, onde os estimadores são menos sensíveis a outliers.

2. **Conexão com a Teoria de Entropia:**

A relação entre medidas de divergência e entropia foi destacada, sugerindo que ambas as abordagens podem ser complementares em análises de séries temporais.

---

#### Dia 5: Implementação Prática das Medidas de Divergência

**Data:** 14 de Outubro de 2024

Hoje comecei a implementar algumas funções básicas para calcular divergências, como a divergência de Kullback-Leibler e divergências quadráticas.

---

```
```python
import numpy as np

def kl_divergence(p, q):
    return np.sum(p * np.log(p / q))
```

## Exemplo de cálculo de divergência

```
p = np.array([0.4, 0.6])
q = np.array([0.5, 0.5])
print(f"Divergência de KL: {kl_divergence(p, q)}")
```

---

#### Dia 6: Discussão Teórica sobre Robustez

**Data:** 15 de Outubro de 2024

A robustez na inferência estatística foi o foco do dia. Leandro Pardo argumenta que métodos baseados em divergência são mais robustos contra violações de suposições modelísticas e outliers.

---

#### Dia 7: Exercícios sobre Inferência por Divergência

**Data:** 17 de Outubro de 2024

Hoje dediquei tempo a resolver exercícios do livro de Pardo, com foco em derivar estimadores baseados na minimização de divergências.

---

#### Dia 8: Aplicação da Inferência

**Data:** 19 de Outubro de 2024

Explorei como inferências baseadas em distância mínima e divergência podem ser aplicadas, complementando a análise de entropia.

---

### **Dia 10: Reflexões e Próximos Passos**

**Data:** 21 de Outubro de 2024

Percebo que o uso de divergências são variados mas não identifico ainda as métricas que os tornam comparáveis.