Cálculo da Função de Correlação Angular de Dois Pontos com o Estimador de Landy-Szalay

Observatório Nacional

July 28, 2025

Introdução

A distribuição das galáxias no Universo não é aleatória: elas tendem a se agrupar em grandes estruturas como aglomerados e filamentos. Uma ferramenta estatística muito utilizada para quantificar essa distribuição é a função de correlação angular de dois pontos, $\omega(\theta)$, que mede o excesso (ou falta) de pares de galáxias a uma certa separação angular θ, em comparação com um catálogo completamente aleatório. Nesta atividade, vamos usar um dos estimadores mais comuns e robustos para calcular essa função: o estimador de Landy-Szalay. Ele combina contagens de pares em catálogos de dados reais e catálogos aleatórios, com a seguinte fórmula:

$$\omega(\theta) = \frac{DD(\theta) - 2DR(\theta) + RR(\theta)}{RR(\theta)}$$

onde $DD(\theta)$ é o número de pares de objetos no catálogo de dados separados por um ângulo θ, $RR(\theta)$ é o número de pares no catálogo aleatório e $DR(\theta)$ é o número de pares entre o catálogo de dados e o aleatório.

Objetivos da atividade

- Familiarizar-se com o conceito de função de correlação angular de 2 pontos;
- Implementar o estimador de Landy-Szalay em um conjunto realista de dados;
- Reproduzir o gráfico de $\omega(\theta)$ semelhante aos da literatura.

Metodologia

Nesta atividade, você deve:

1. Carregar um catálogo de objetos celestes contendo coordenadas (Ra, Dec);
2. Construir histogramas de separações angulares entre pares de objetos do catálogo (para DD);
3. Fazer o mesmo para pares entre catálogos aleatórios (RR) e entre dados e random (DR);

4. Usar o estimador de Landy-Szalay para calcular $\omega(\theta)$;
5. Plotar $\omega(\theta)$ em função de θ.

Para facilitar o trabalho, já disponibilizamos:

- Um catálogo de dados reais (arquivo `ALFALFA_RA_113-155_DEC_0-18_2081objects.csv`);
- Dez catálogos aleatórios (arquivos `random_1.dat`, `random_2.dat`, ..., `random_10.dat`);
- Um gráfico de exemplo com o resultado esperado.

Normalização das contagens

Ao calcular os pares $DD(\theta)$, $DR(\theta)$ e $RR(\theta)$, é necessário normalizá-los para que as contagens sejam comparáveis entre si, independentemente do número de objetos em cada catálogo. Isso é feito dividindo cada contagem pelo número total de pares possíveis em cada caso:

$$DD(\theta) = \frac{\text{número de pares no catálogo de dados com separação } \theta}{N_D(N_D - 1)/2},$$

$$RR(\theta) = \frac{\text{número de pares no catálogo aleatório com separação } \theta}{N_R(N_R - 1)/2},$$

$$DR(\theta) = \frac{\text{número de pares entre dados e random com separação } \theta}{N_D \cdot N_R},$$

onde:

- N_D é o número de objetos no catálogo de dados;
- N_R é o número de objetos em um catálogo aleatório.

Descrição dos dados

Os arquivos estão em formato CSV, com colunas separadas por tabulação horizontal (`\t`). Cada linha representa um objeto.

- **Catálogo de dados:** `ALFALFA_RA_113-155_DEC_0-18_2081objects.csv`
  - **Coluna 2:** Nome do objeto;
  - **Coluna 3:** RAdeg_HI (ascensão reta) em graus;
  - **Coluna 4:** DECdeg_HI (declinação) em graus;
  - As demais colunas envolvem o fluxo dos objetos, razões sinal-ruído etc., e podem ser ignoradas para a presente atividade.
- **Catálogos aleatórios:** `random_X.dat` ($X = 0$ a 9)
  - Mesmo formato do catálogo de dados: Ra e Dec em graus

– Mesma quantidade de dados que o catálogo original, por questões de custo computacional. Na pasta de catálogos aleatórios, encontra-se um arquivo `.py` contendo o código utilizado para gerar esses catálogos, o qual pode ser modificado caso se deseje aumentar o número de objetos nos randoms para testes adicionais.;

Resultado esperado

Ao final da atividade, você deve obter um gráfico semelhante ao presente na pasta do Drive `correlation_function.png`, que mostra a função $\omega(\theta)$ em função da separação angular θ. Um valor positivo de ω indica excesso de correlação (aglomeração), enquanto um valor negativo sugere anticorrelação.