Mathematical Optimization Algorithms

Lab 2: Statistics





Comprehensive Study of World Happiness

Scenario:

 You are a consultant working for the United Nations tasked with preparing a comprehensive study on global happiness. Utilizing a dataset that includes various countries and the following attributes social_support, freedom, corruption, generosity, GDP per capita, life expectancy, and happiness score you are to analyze the factors that influence happiness around the world.

Objective:

- Based on the given dataset, formulate and answer several research questions using Python and relevant libraries such as pandas, numpy, matplotlib, and scipy. Your analysis should include:
- Descriptive statistics of key variables.
- Probability calculations related to the attributes.
- Inferential statistics to draw conclusions about the global population.
- Hypothesis testing to evaluate assumptions about the data.



Comprehensive Study of World Happiness

Statistics

- 1. Calculate the mean, median, and standard deviation of the happiness_score across all countries.
- 2. Find the top 5 countries with the highest gdp_per_cap and plot their happiness_score in a bar chart.
- 3. Compute the correlation matrix for all numerical variables in the dataset.
- 4. Plot a histogram of life_exp and analyze its distribution.

Probability

- 5. If a country is selected at random, what is the probability that its social_support score is greater than 50?
- 6. Assuming that the life_exp values are normally distributed, what is the probability that a randomly selected country has a life_exp greater than 80 years?

Inference

7. Perform a linear regression analysis to predict happiness_score based on gdp_per_cap.

Hypothesis Testing

8. Is there a significant difference in life_exp between countries with social_support above and below the median? Perform an appropriate hypothesis test.

Mathematical Optimization Algorithms

Lab 2: Statistics



