📊 COMPLETE STATISTICS SYLLABUS FOR DATA SCIENCE

A Structured Learning Path for Aspiring Data Scientists

🎯 Course Overview

Complete Statistics Syllabus: This document outlines the comprehensive 15-section statistics curriculum designed specifically for data science aspirants. Each section builds upon previous knowledge to create a strong statistical foundation.

1. Types of Data

Qualitative (Categorical)

Nominal

Ordinal

Quantitative (Numerical)

Discrete

Continuous

2. Types of Statistics

Descriptive Statistics

Inferential Statistics

3. Descriptive Statistics

Measures of Central Tendency

- Mean
- Median
- Mode

Measures of Dispersion

- Range
- Variance
- Standard Deviation
- Interquartile Range (IQR)

Measures of Shape

- Skewness
- Kurtosis

Percentiles and Quartiles

Frequency Distribution

Data Visualization Concepts

- Histogram
- Box Plot
- Bar Chart
- Pie Chart
- Line Chart

4. Inferential Statistics

Population vs Sample

Sampling Techniques

- Random Sampling
- Stratified Sampling
- Systematic Sampling
- Cluster Sampling

Sampling Distribution

Central Limit Theorem

5. Probability Theory

Basic Probability Concepts

Types of Probability

Classical Probability

Empirical Probability

Subjective Probability

Rules of Probability

Addition Rule

Multiplication Rule

Conditional Probability

Bayes' Theorem

6. Random Variables

Discrete Random Variables

Continuous Random Variables

Probability Mass Function (PMF)

Probability Density Function (PDF)

Cumulative Distribution Function (CDF)

Expected Value

Variance of Random Variables

7. Probability Distributions

Discrete Distributions

Bernoulli Distribution

Binomial Distribution

Poisson Distribution

Continuous Distributions

Uniform Distribution

Normal (Gaussian) Distribution

Exponential Distribution

Log-Normal Distribution

8. Hypothesis Testing

Null Hypothesis

Alternative Hypothesis

Type I Error

Type II Error

Significance Level (α)

p-value

Test Statistic

Critical Value

9. Statistical Tests

Z-test

t-test

- One-sample t-test
- Two-sample t-test
- Paired t-test

Chi-Square Test

ANOVA

Non-Parametric Tests

Mann-Whitney U Test

Wilcoxon Signed-Rank Test

Kruskal-Wallis Test

10. Confidence Intervals

Confidence Level

Margin of Error

Interval Estimation for Mean

Interval Estimation for Proportion

11. Correlation and Covariance

Covariance

Correlation Coefficient

Pearson Correlation

Spearman Rank Correlation

12. Regression Analysis

Simple Linear Regression

Multiple Linear Regression

Assumptions of Regression

Coefficient Interpretation

R-squared

Adjusted R-squared

Residual Analysis

13. Statistical Assumptions

Normality

Linearity

Independence

Homoscedasticity

Multicollinearity

14. Exploratory Data Analysis (EDA)

Data Cleaning Concepts

Missing Value Handling

Outlier Detection

Data Transformation

Data Scaling

Feature Distribution Analysis

15. Advanced Statistical Concepts (Data Science Focused)

Law of Large Numbers

Bias and Variance

Overfitting and Underfitting

Bootstrap Sampling

Cross-Validation

A/B Testing

Time Series Components

Trend

Seasonality

Noise

Complete Statistics Syllabus for Data Science
15 Comprehensive Sections - Structured Learning Path

Data Science with Vamsi