

Introduction to Applied Statistics

Introduction to Applied Statistics

Key concepts, workflow, and practical tools for turning data into insight.

1. What Is Applied Statistics?

Aspect	Description
Definition	The use of statistical methods to collect, analyze, interpret, and present data for real-world problems.
Goal	Turn raw data into actionable information for decision-making in fields such as business, health, engineering, social sciences, and public policy.
Contrast with Theoretical Statistics	Theoretical statistics focuses on developing new methods and proving their properties; applied statistics emphasizes implementation, model selection, and communication of results.

2. The Applied Statistics Workflow

- Problem Definition** – Clarify the research question, objectives, and success criteria.
 - Data Collection** – Choose a sampling plan, design experiments, or retrieve existing datasets.
 - Data Preparation** – Clean, transform, and explore data (missing values, outliers, variable types).
 - Exploratory Data Analysis (EDA)** – Visualize distributions, relationships, and spot patterns.
 - Model Building** – Select appropriate statistical models (regression, classification, time series, etc.).
 - Model Assessment** – Check assumptions, evaluate fit, and validate with cross-validation or hold-out tests.
 - Inference & Interpretation** – Estimate parameters, test hypotheses, construct confidence intervals, and translate findings into plain language.
 - Communication** – Create tables, graphs, and reports tailored to the audience; highlight limitations and recommendations.
-

Introduction to Applied Statistics

3. Core Statistical Concepts

Concept	Why It Matters	Typical Tools
Population vs. Sample	Inferences are made about a larger group using a subset.	Simple random sampling, stratified sampling
Variable Types	Determines which analyses are appropriate.	Categorical, ordinal, interval, ratio
Descriptive Statistics	Summarize central tendency, spread, and shape.	Mean, median, mode, variance, standard deviation, IQR, histograms, boxplots
Probability Distributions	Model randomness and underpin inferential methods.	Normal, binomial, Poisson, exponential, t? distribution
Estimation	Quantify unknown parameters from data.	Point estimates, confidence intervals
Hypothesis Testing	Assess evidence for/against a claim.	p?values, type I/II errors, power, common tests (t, ? ² , ANOVA)
Regression & Correlation	Model relationships between variables.	Simple/Multiple linear regression, logistic regression, Pearson/Spearman correlation
ANOVA & MANOVA	Compare means across multiple groups.	One?way, two?way ANOVA, post?hoc tests
Non?parametric Methods	Useful when distributional assumptions fail.	Wilcoxon, Kruskal?Wallis, Mann?Whitney
Resampling Techniques	Estimate sampling variability without strong parametric assumptions.	Bootstrap, permutation tests
Time?Series & Forecasting	Analyze data indexed in time order.	ARIMA, exponential smoothing, seasonal decomposition
Multivariate Analysis	Examine several variables simultaneously.	Principal Component Analysis (PCA), cluster analysis, discriminant analysis

4. Common Data Types & Sources

Domain	Typical Data	Collection Methods
Page 2 Back to Ynet StudyHub		

Introduction to Applied Statistics

Business Sales, customer churn, web traffic Transaction logs, surveys, APIs

Health & Medicine Patient outcomes, biomarker levels Clinical trials, electronic health records

Engineering