Interdisciplinary Symposia Series 2025

MTTF
MathTech Thinking Foundation
A Ministry of Corporate Affairs, Government of INDIA Approved NGO

Hands-on Workshop on

Biostatistics for Data Analysis in Biology and Medicine Using R Programming

(Speaker)
Dr. Mehar Chand
Baba Farid College of Engineering and Technology, BTI and President, MTTF, Fazilka, India

(Coordinator)
Ms. Muskan Gupta
Project Associate, Department of Pharmacology, AIIMS, Bathinda, India

(Coordinator)
Dr. Aparna Bararia (PT)
Department of Biomedical and Allied Health Sciences, Baba Farid College of Engineering and Technology, BTI

The Role of Biostatistics in Research

Experiment Design
Planning and structuring research studies

Data Analysis
Processing and interpreting research data

Conclusion Drawing
Deriving meaningful insights from research

Genetics
Studying hereditary traits and genetic variations

Epidemiology
Analyzing patterns and causes of diseases

Clinical Trials
Evaluating the effectiveness of medical treatments

Public Health
Improving community health through research

About the workshop

The Hands-on Workshop on Biostatistics for Data Analysis in Biology and Medicine Using R Programming offers practical training in applying statistical methods to real biological and medical datasets. Participants will learn data handling, hypothesis testing, regression, survival analysis, and visualization using R. This workshop is ideal for students, researchers, and healthcare professionals aiming to enhance their analytical and research skills.

Who can participate?

- Students from biology, medicine, biotechnology, and allied fields
- Research scholars and academicians involved in life science studies
- Healthcare and pharmaceutical professionals handling research data
- Anyone interested in learning biostatistical data analysis using R programming

Limited SEATS

Registration Fee:

National Participant: 500 INR

International Participant: 10 USD

Event Date: October 25-31, 2025

Duration: 19:00-21:00 IST

Mode: Virtual (Online)

Scan QR code for registration

MATHTECH THINKING FOUNDATION

Corporate office: # 4, Fazilka-Abohar, Road Near Bus Stand, Fazilka-152123 (Punjab), INDIA

www.mttf.in +91-828-809-0404 contactus@mttf.in

Overview

The Hands-on Workshop on **Biostatistics for Data Analysis in Biology and Medicine Using R Programming** is designed to provide participants with practical skills to apply statistical methods in biological, biomedical, and health-related research. With the increasing availability of complex datasets in life sciences, the ability to analyze, interpret, and visualize data using reliable tools is an essential skill for students, researchers, and professionals.

This workshop integrates biostatistical concepts with hands-on training in R programming, enabling participants to bridge the gap between theory and practice. Through interactive sessions, case studies, and guided exercises, attendees will learn how to manage datasets, perform descriptive and inferential statistics, build predictive models, conduct survival and multivariate analyses, and create effective data visualizations.

By the end of the workshop, participants will be able to:

- Understand the role and importance of biostatistics in biology, medicine, and healthcare research.
- Apply a wide range of statistical methods using R programming.
- Analyze biological, clinical, and epidemiological datasets effectively.
- Generate meaningful visualizations and reports for scientific communication.
- Design reproducible workflows for research projects.

This program is ideal for students, academicians, researchers, healthcare professionals, and data enthusiasts who want to strengthen their analytical capabilities and apply biostatistics to solve real-world biological and medical problems.

Objectives of the Workshop

- To introduce fundamental concepts of biostatistics and their applications in biological and medical research.
- To familiarize participants with data handling, visualization, and analysis using R programming.
- To enable participants to perform statistical tests and interpret results effectively.
- To provide hands-on experience in analyzing real biological and healthcare datasets.
- To promote the use of open-source tools for reproducible and data-driven research.

Learning Outcomes

- To introduce fundamental concepts of biostatistics and their applications in biological and medical research.
- To familiarize participants with data handling, visualization, and analysis using R programming.
- To enable participants to perform statistical tests and interpret results effectively.
- To provide hands-on experience in analyzing real biological and healthcare datasets.
- To promote the use of open-source tools for reproducible and data-driven research.

Key Highlights

- Interactive sessions combining theory with practical exercises.
- Hands-on training using real biological and healthcare datasets.
- Step-by-step guidance on data visualization and statistical modeling in R.
- Expert-led lectures and live demonstrations.
- E-certificates for all active participants.
- Ideal for students, researchers, and professionals in life sciences and medicine.

Why R Programming for Biostatistics

- Open-source and free: Accessible to everyone without licensing costs.
- Powerful statistical tools: Offers a vast range of built-in functions for biostatistical analysis.
- Excellent data visualization: Creates high-quality and customizable graphs for research.
- Reproducible research: Ensures transparency and repeatability in data analysis.
- Widely used in academia and healthcare research: Makes it a valuable skill for modern biologists and medical scientists.

Applications of R Programming

- **Clinical Research:** Analyzing patient data, treatment outcomes, and clinical trials.
- **Epidemiology:** Studying disease patterns, prevalence, and risk factors.
- **Genomics & Bioinformatics:** Gene expression analysis, sequencing data, and biomarker discovery.
- **Public Health:** Health surveys, population studies, and intervention assessments.
- **Pharmaceutical Research:** Drug efficacy, toxicity studies, and pharmacovigilance.
- **Experimental Biology:** Lab experiments, growth studies, and ecological data analysis.

Useful R Packages for Biostatistics

Data Handling & Manipulation

- tidyverse – Data cleaning, transformation, and visualization
- dplyr / tidyr – Organize, filter, and reshape datasets

Descriptive & Inferential Statistics

- MASS – Applied statistics and biomedical datasets
- boot – Bootstrapping and resampling methods
- TH.data – Companion datasets for practice
- psych – Descriptive statistics and psychometric analysis

Regression & Predictive Modeling

- caret – Machine learning and predictive modeling
- glmnet – Regularized regression models
- lmtest – Diagnostic testing for regression models

Survival & Clinical Analysis

- survival – Kaplan-Meier, Cox regression, time-to-event analysis
- survminer – Enhanced visualization for survival curves

Data Visualization

- ggplot2 – Advanced and customizable plots
- pheatmap – Heatmaps for gene expression or multivariate data
- corrplot – Correlation matrices visualization

Multivariate & Clustering

- cluster – Clustering methods (hierarchical, k-means)
- factoextra – Visualizing PCA and clustering results

Bioinformatics / Genomics

- Bioconductor packages (e.g., airway, edgeR, DESeq2) – RNA-seq and genomic analysis

Reproducibility & Reporting

- rmarkdown – Combine code, analysis, and plots into reproducible reports

Day-Wise Content Plan

Day 1: Introduction to Biostatistics & R Programming

- Importance of Biostatistics in Biological Research
- Overview of R & RStudio: Installation, Interface, and Basic Commands
- Data Types, Data Frames, and Importing Biological Data (CSV, Excel, TXT)
- Descriptive Statistics: Mean, Median, Mode, Variance, SD
- Hands-on in R: Data manipulation & visualization (histograms, boxplots, barplots)

Day 2: Probability & Statistical Distributions

- Concepts of Probability in Biological Studies
- Probability Distributions: Normal, Binomial, Poisson
- Sampling Methods & Central Limit Theorem
- Hands-on in R: Generating distributions, probability density functions, random sampling
- Biological Example: Modeling disease occurrence / mutation probabilities

Day 3: Statistical Inference & Hypothesis Testing

- Population vs. Sample, Confidence Intervals
- Hypothesis Testing Framework: p-values, Type I & II errors
- Parametric Tests: t-test (independent, paired), ANOVA
- Hands-on in R: Applying t-tests & ANOVA on biological datasets
- Case Study: Gene expression data / clinical trial sample analysis

Day 4: Correlation & Regression Analysis

- Correlation (Pearson, Spearman) & Interpretation in Biology
- Simple Linear Regression
- Multiple Linear Regression Models
- Hands-on in R: Correlation matrices, regression models, diagnostic plots
- Biological Example: Effect of environmental factors on plant/animal growth

Day 5: Non-Parametric & Advanced Methods

- When to use Non-Parametric Tests
- Mann–Whitney U test, Wilcoxon Signed-Rank, Kruskal–Wallis test
- Chi-Square test for categorical data
- Hands-on in R: Executing non-parametric tests
- Biological Example: Comparing treatment groups with small sample sizes

Day 6: Survival Analysis & Multivariate Statistics

- Introduction to Survival Data in Medicine (time-to-event data)
- Kaplan–Meier Curves, Log-rank Test, Cox Proportional Hazards Model
- Multivariate Analysis: Principal Component Analysis (PCA), Cluster Analysis
- Hands-on in R: Survival package, survival curves, PCA & clustering with biological datasets
- Case Study: Clinical survival data & genetic expression profiling

Day 7: Project Work, Reporting & Visualization

- Designing a Biostatistical Study in Biology/Medicine
- Reproducible Research with RMarkdown
- Data Visualization: ggplot2 for effective plots & graphics
- Hands-on Project: Analyzing a real biological dataset (group activity)
- Presentation of Results & Scientific Report Writing
- Workshop Wrap-Up: Feedback, Q&A, Certificates