

Course Syllabus

BIOENG-210
Prof. Gioele La Manno

Edition 2025

Syllabus Topics

I. Probability Random Variables

Probability Foundations

- Random variables: definition and types
- Functions describing random variables

Expectation and Moments

- Definition and properties of expectation
- Central moments and variance

Parametric Distributions

- Normal and Log-Normal distributions
- Gamma, Beta and Dirichlet distributions
- Count distributions (Poisson, Neg. Binomial)

Operations on Random Variables

- Convolution and mixture distributions
- Transformation of random variables

II. Statistical Inference

Maximum Likelihood Estimation

- Likelihood function and log-likelihood
- Finding maximum likelihood estimates
- Bias in MLE

Measuring Association

- Covariance and correlation
- Geometric interpretation
- Spearman's rank correlation

Joint Probability Distributions

- Marginal and conditional distributions
- Independence concepts
- Bivariate normal distribution

Estimation of Joint Distributions

- Parametric and non-parametric estimation
- Mutual information

III. Statistical Testing

Foundations of Statistical Testing

- Test statistics and probability calculations
- Central limit theorem

Testing Scenarios

- One/two-sample and paired tests
- Likelihood ratio tests

Multiple Testing

- Family-wise error rate (FWER)
- False discovery rate (FDR)
- Benjamini-Hochberg procedure

IV. Regression Analysis

Simple Linear Regression

- Probabilistic perspective
- Parameter estimation
- Inference and confidence intervals

Multiple Linear Regression

- Design matrix and parameter estimation
- Interpretation of coefficients

Model Comparison

- Assessing model fit (R^2 and adjusted R^2)
- Likelihood ratio tests
- ANOVA in regression framework

V. Generalized Linear Models

- Link functions
- Poisson regression for count data
- Logistic regression for binary outcomes
- ROC curves and model evaluation

VI. Regularization and Bayes

Beyond Maximum Likelihood

- Overfitting and generalization
- Cross-validation

Shrinkage Methods

- Ridge regression
- Bias-variance trade-off
- Choosing regularization parameters

Bayesian Approaches

- Maximum a posteriori estimation
- Lasso regression for feature selection

VII. Multivariate Analysis

Multivariate Normal Distribution

- Covariance matrix
- Geometric properties

High-Dimensional Data Analysis

- Challenges in estimation
- Eigenvalue analysis

Principal Component Analysis

- Eigendecomposition and SVD
- Dimensionality reduction
- Interpretation of principal components

VIII. Clustering Methods

K-Means Clustering

- Algorithm and properties
- Choosing optimal K

Hierarchical Clustering

- Agglomerative methods
- Building cluster hierarchies

IX. Resampling Methods

- Jackknife for standard errors
- Permutation tests
- Bootstrap for statistical inference
- Monte Carlo simulation principles