BSDS: Statistics II

First Course Handout

Instructor in charge: Minerva Mukhopadhyay (minervam@isical.ac.in)

Tutors: **1) Kolkata:** Soutik Halder (haldersoutik2015@gmail.com)

2) Delhi: Anirban Mondal (anirban.mondal13@gmail.com)

3) Bangalore: Debanjana Datta (rs_asu2401@isibang.ac.in)

Lecture hours: TTh 10:30AM-12 noon **Tutorial:** Th 4-5PM

Course webpage: https://shorturl.at/bW9cu

Objective: This course explores key concepts in classical statistics, providing students with a comprehensive introduction to statistical inference. Topics include frequentist estimation and hypothesis testing, with an emphasis on solving problems both mathematically and through computer simulations. Real-life examples will be used to illustrate the application of statistical methods, and the effectiveness of these procedures will be evaluated through numerical simulations.

References:

1. Statistics by D. A. Freedman, R. Pisani and R. Purves.
2. The Art of Statistics: How to Learn from Data by D. Spiegelhalter.
3. All of Statistics: A Concise Course in Statistical Inference by L. Wasserman.
4. Probability and Statistics by M. H. DeGroot and M. J. Schervish.
5. Mathematical Statistics with Applications by D. Wackerly, W. Mendenhall, R. L. Scheaffer
6. Probability and Statistics by Siva Athreya, Deepayan Sarkar, and Steve Tanner [LINK]

Course Content:

1. Population vs. sample. Empirical distribution. Parametric statistical models. Need for inference on a parameter.
2. Point estimation: Concept of bias, variance, mean squared error (MSE), relative efficiency.
3. Estimation of parameters by method of moments and maximum likelihood method.
4. Sampling distributions derived from normal populations: Chi-square, t and F distributions. (Discussion of properties without derivation).
5. Introduction to asymptotic inference: Informal discussion of Weak Law of Large Numbers and Central Limit Theorem. Consistency of estimators. Approximating sampling distributions through CLT.
6. Numerical computation of sampling distributions based on independent samples from an arbitrary population. Comparison with CLT-based approximations.

7. Hypothesis testing: Null and alternative hypotheses, simple and composite hypotheses, Type I and Type II errors, level and power of a test, p-value. Neyman-Pearson Lemma, most powerful tests, unbiased tests.
8. Exact and large sample tests for binomial proportion and mean of a normal distribution (one sample case). One-sided and two-sided alternatives.
9. Exact (pooled-t) and large sample tests for the equality of two normal means. Paired t-test. Test for equality of variances of two normal populations. Test for the equality of two binomial proportions.
10. Interval estimation: Construction using pivotal quantities and critical regions. Exact and large sample confidence intervals. Confidence intervals for binomial, normal and Poisson parameters (one sample case).
11. Large sample confidence intervals for the difference of two binomial proportions. Exact and large sample confidence intervals for the difference of means of two normal populations.
12. Sample size determination in tests of hypotheses for ensuring a specified power, and in constructing confidence intervals for ensuring a specified interval width.

Grading and Evaluation:

(i) There will be 2-3 **quizzes**, a mid-semester and an end-semester exams.

(ii) Questions will be of objective and/or descriptive type.

(iii) Marking scheme is *absolute*. The weights of different evaluations are as follows: Quizzes: 30% on total score in quizzes (*)
Midterm: 30%
Endterm: 40%
(*) A student's total score in quizzes will be the sum (weighted sum if the total marks of the quizzes are different) of their two highest-scoring quizzes (in cases three quizzes are conducted).

(iv) If a student misses a quiz, then s/he will be evaluated based on the other quizzes. If s/he misses more than one quizzes, then s/he will get zero marks in the missed quizzes.

(v) The answer scripts after evaluation will be uploaded in the GradeScope.

General Guidelines:

(i) Students are expected to spend 12 hours per week for good performance in this course.

(ii) If a student found to follow unfair means (in any of the exams), then irrespective of the circumstances his/her exam will be cancelled (with no supplementary exam).