Columbia University
MAILMAN SCHOOL
OF PUBLIC HEALTH

Statistical Inference P8109 (3 points)

COURSE SYLLABUS
Spring 2022

CLASS SESSION

Course	Day	Time
P8109	Mon	1:00-3:50pm

The first (Jan 24) class will be done on Zoom:
https://columbiacuimc.zoom.us/j/91453599543
Passcode: P8109-001

All subsequent classes will be done in:
Alumni Auditorium Located at 650 W. 168th St., (accessible via the William Black Building)

INSTRUCTOR

Dr Prakash Gorroochurn, Associate Professor
Office Phone: (212)342-1263
E-mail: pg2113@columbia.edu
Office: Allen Rosenfeld (Old PI) Building, R620
Homepage: http://www.columbia.edu/~pg2113/

CONSULTATION

Thursday: 1:30 pm -2:30 pm either in R620 or on zoom:
https://us02web.zoom.us/j/5389716351?pwd=Tmh1b3VFSW5ldHVvQzZBVTJBQ3Budz09
passcode: 4evsiN

TA's

- Chai, Zilan (zc2326@cumc.columbia.edu)

Monday 5:30-6:30pm via Zoom (https://columbiacuimc.zoom.us/j/95556324990)

- Gacheru, Margaret (mg3861@cumc.columbia.edu)

Wednesday 11am - 12pm via Zoom

(https://zoom.us/j/96491741265?pwd=TXR0b1JUMFNEMVI5cVB0RHZVbzVTUT09)

- Tumasian, Robert A. (rat2134@cumc.columbia.edu)

Tuesday from 11am-12pm via Zoom (link: https://columbiacuimc.zoom.us/j/2383237115)

- Wang, Siquan (sw3442@cumc.columbia.edu)

Friday 3 pm - 4 pm via Zoom (https://columbiacuimc.zoom.us/j/9015192177).

- Yao, Yujing (yy2725@cumc.columbia.edu)

Thursday 11am-12pm via Zoom (https://columbiacuimc.zoom.us/j/91274968043).

COURSE DESCRIPTION

The course aims to present the fundamental principles behind statistical inference. The course starts by presenting major topics in distribution theory. The subject of statistical estimation is then covered with the associated topics of likelihood, properties of estimators, Cramér-Rao inequality, sufficiency, and completeness. The subject of hypothesis testing is then covered, with associated topics like power, Neyman-Pearson lemma, uniformly most powerful tests, and unbiased tests.

PREREQUISITES

P8104 and a working knowledge of calculus and linear algebra.

ASSESSMENT AND GRADING POLICY

Student grades will be based on the following:

- Attendance of at least 9 of Sessions 1-11, worth 5% of the final grade.
- 5 assignments to be completed at home individually, worth in all 30% of the final grade. *ALL ASSIGNMENTS SHOULD BE SUBMITTED ONLINE.*
- One mid-term examination to be done in the class on March 14, worth 15% of the final grade.
- One final examination based on the entire course on April 25, to be done in the class, worth 50% of the final grade.

The final course grade will be determined using the School's letter grade system. Grades are: **A, B, C**, with **+** and **-** as applicable. Grades are defined as follows:

A+	Reserved for highly exceptional achievement.
A	Excellent. Outstanding achievement.
A-	Excellent work, close to outstanding.
B+	Very good. Solid achievement expected of most graduate students.
B	Good. Acceptable achievement.

B- Acceptable achievement, but below what is generally expected of graduate students.
C+ Fair achievement, above minimally acceptable level.
C Fair achievement, but only minimally acceptable.
C- Very low performance.
F Failure. Course usually may not be repeated unless it is a required course.

TEXTBOOKS

- Required: Introduction to the Theory of Statistics (1974) by Mood A.M., Graybill F.A. & Boes D.C., 3rd edition. McGraw-Hill (ISBN: 978-0070428645)
- Recommended: Statistical Inference (2001) by Casella G. & Berger R.L., 2nd edition. Duxbury Press (ISBN: 978-0534243128)

Students who wish to review basic concepts of probability and statistics, and who need additional practice questions should also consult:

- Mathematical Statistics with Applications (2008) by Wackerly D., Mendenhall W & Scheaffer R.L., 7th edition (ISBN: 9780495110811)

Students who are interested in the historical aspects of inference may consult my book:

- Classic Topics on the History of Modern Mathematical Statistics: From Laplace to Modern Recent Times (2016) by Gorroochurn P. Wiley (ISBN: 978-1119127925)

COURSE SCHEDULE

N.B. Brackets Refer to Chapters and Sections of the Required Textbook by Mood et al.

SESSION 1 (Jan 24) - NORMAL AND ASSOCIATED DISTRIBUTIONS
Learning Objectives

Understand:

- **The importance of the normal distribution (VI-4.1)**
- **The nature of the $\chi 2$-, t-, and F-distributions (VI-4.3, VI-4.4, VI-4.5)**

Prove:

- **The Central Limit Theorem (CLT) (VI-3.3)**

SESSION 1 only will conducted via Zoom:
https://columbiacuimc.zoom.us/j/91453599543
Passcode: P8109-001
SESSION 2 onwards will be held in the Alumni Auditorium.

SESSION 2 (Jan 31) - NORMAL AND ASSOCIATED DISTRIBUTIONS
Learning Objectives

Understand:

- **The distributions that arise by sampling from a normal distribution (VI-4)**

- The distributions of the sample mean and sample variance (VI-4.2, VI-4.3)

Prove:

- The independence between the sample mean and sample variance when sampling from a normal distribution (VI-4.3)

SESSION 3 (Feb 7) - NORMAL AND ASSOCIATED DISTRIBUTIONS
Learning Objectives

Understand:
- The importance of the exponential family (VII-4.4)

Use:
- Lévy's Continuity theorem
- Slutsky's Theorem
- Cramér's delta theorem

SESSION 4 (Feb 14) - ESTIMATION
Learning Objectives

Understand:
- The Importance of Statistical Models
- The concept of likelihood (VII-2.2)
- The properties of unbiasedness and consistency (VII-3.2 VII-3.3)
- The concept of uniform minimum variance unbiased estimators (UMVUE) (VII-5)

Derive:
- Maximum likelihood estimators (MLEs) (VII-2.2)

SESSION 5 (Feb 21) - ESTIMATION
Learning Objectives

Understand:
- The concept of Fisher information
- The Cramér-Rao Inequality (VII-5.1)
- Asymptotic properties of maximum likelihood estimators (MLEs) (VII-9)

Prove:
- The Cramér-Rao Inequality (VII-5.1)

SESSION 6 (Feb 28) - ESTIMATION
Learning Objectives

Understand:
- The Concept of Sufficiency (VII-4)
- The Fisher-Neyman Factorization Theorem (VII-4.2)

- The Rao-Blackwell Theorem (VII-5.2)
- The concept of minimal sufficiency (VII-4.3)

Prove:

- The Fisher-Neyman Factorization Theorem (VII-4.2)
- The Rao-Blackwell Theorem (VII-5.2)

<u>MIDTERM EXAM (March 14)</u>

Midterm exam in class. Covers Sessions 1-6

<u>SESSION 7 (Mar 21) - ESTIMATION</u>
Learning Objectives

Understand:

- The Concept of Completeness (VII-5.2)
- The Lehmann-Scheffé theorem (VII-5.2)

Prove:

- The Lehmann-Scheffé theorem (VII-5.2)

<u>SESSION 8 (March 28) – CONFIDENCE LEVELS AND TESTS</u>
Learning Objectives

Understand:

- The interpretation of confidence intervals (VIII-2)
- The nature of statistical hypotheses (IX-1)
- One-sided and two-sided hypotheses (IX-4.1)
- The concept of the size of a statistical test (IX-1)
- The concept of power function (IX-1)
- The concept of pivot (VIII-2.3)

<u>SESSION 9 (April 4) - CONFIDENCE LEVELS AND TESTS</u>
Learning Objectives

Understand:

- The Neyman-Pearson (NP) lemma (IX-2.2)
- The generalised likelihood ratio test (IX-3.1)

Prove:

- The Neyman-Pearson (NP) lemma (IX-2.2)

Derive:

- Most powerful tests (IX-2.2)

SESSION 10 (April 11) - CONFIDENCE LEVELS AND TESTS
Learning Objectives

Understand:

- **The concept of uniformly most powerful (UMP) tests (IX-3.2)**
- **The property of monotone likelihood ratio (MLR) (IX-3.2)**
- **The concept of unbiased tests (IX-3.3)**

SESSION 11 (April 18) - BAYESIAN ESTIMATION

Learning Objectives

Understand:

- **The nature of posterior Bayes estimators (VII-7.1)**
- **The notions of loss and risk functions (VII-7.2)**

Construct:

- **Bayes estimators (VII-7.1)**
- **Minimax estimators (VII-7.3)**

April 25 - FINAL EXAM