MATH 5010-Foundations of Statistical Theory & Probability – Fall 2025

Instructor: He Wang
Class time and room: see Canvas
Office Hours: see Canvas
Email: he.wang@northeastern.edu

Recommended textbooks
More course material will be available on Canvas. Recommended Textbooks:

- **Statistical Inference, by Casella, George, Berger, Roger L, 2nd edition (Main Reference)**
- Introduction to Probability. C.M. Grinstead and J.L. Snell. American Mathematical Society,2012
- Probability and Statistics, by DeGroot and Schervish
- Introduction to Mathematical Statistics (Hogg, McKean, Craig)

Prerequisite: Basic knowledge (undergraduate level) about linear algebra, multivariable calculus, probability and statistics are required.

Course Description
Offers a comprehensive foundational overview of the probability theory and mathematical statistics critical to various methodologies needed to perform and interpret a statistical analysis. Beginning with an exploration of probability distributions, the course explores the theoretical justifications around common statistical measures such as means, variances, and expected values. Likelihood and Bayesian theory are introduced as tools to derive parameter estimates and begin to make inferences. Culminates with formal procedures for statistical inference, including hypothesis testing and the construction of confidence intervals and credible intervals, and calculations of power and sample size. All topics will be grounded in theory, with appropriate practice using a programming language and real-world datasets to communicate and visualize data and results.

Program Learning Outcomes (PLOs)

- Perform appropriate classical statistical, causal, machine learning, and AI analyses based on solid understanding of the inferential and predictive modeling assumptions and properties.
- Use statistical theory to derive and interpret appropriate and/or novel models for data analyses.
- Communicate and present statistical designs, models, and results effectively across technical and non-technical audiences.

Course Learning Outcomes

- Demonstrate mathematical skills to perform a statistical analysis

- Interpret statistical analyses using the probability principles
- Use statistical software to evaluate probability distributions and manage data
- Use statistical software (R/Python) to learn basics of a statistical programming
- Use data visualization to make results easy to understand
- Communicate findings to non-statisticians

Course Topics (Teaching order may be different)

Part 1: Probability

1. Introduction, Foundations of Probability theory (CB-**§**1.1-1.3)
2. Random variables and distributions (CB-**§**1.4-1.6)
3. Transformations, Functions of a random variable. (CB-**§**2.1, 4.3)
4. Expectations, Moments, **Moment-generating functions**. (CB-**§**2.2-2.3)
5. Common families of Distributions, exponential families (CB-**§**3)
6. **Conditional Probability**: Conditional Expectation and Conditional Distribution. (CB-**§**4.1-4.2)
7. **Multivariate distributions,** Joint probability distributions, Covariance and correlation; Multinomial distribution; Multivariate Normal. (CB-§4.5-4.6)
8. **Inequalities**: Markov, Chebyshev, Chernoff, and Jensen; special discrete distributions (CB-§3.6, 4.7)

Part 2: Statistical Inference

1. **Random samples, Order statistics**; basic simulation theory: Monte Carlo integration, importance sampling (CB-§5.1-5.4)
2. **Large Sample Theory**: Law of large numbers; Central Limit theorem (CB-§5.5)
3. Sufficient statistics (CB-§6.2)
4. **Point Estimation**: Maximum likelihood estimation (MLE); Method of moments (MoM) and their properties (unbiasedness, consistency, efficiency). (CB-§7)
5. **Interval Estimation**: Confidence Intervals. (CB-§9)
6. **Hypothesis Testing** and power of tests (CB-§8)
7. Analysis of Variance (**ANOVA**) (CB-§11)
8. Bayesian Statistics: Prior and Posterior Distributions, Bayesian Point Estimation, Bayesian Interval Estimation, Bayesian Testing Procedures, Gibbs Sampler

Grading: Students will be evaluated on the following basis.

1. **Homework.** (40%)
2. **Computer labs (**20%**)**
3. **Attendance and Class discussion participation (**5%)
4. **Two Midterms.** (20%)
5. **Final Exam**. (15%)

Software: Python will be used throughout the course. Students should be prepared to use Python for lab assignments and for the final project. (It is also ok if student prefer to R). Familiar with any one computer language is required, e.g., Python or Mathlab or R.

Late submission policy. Late submission within a week of any assignment without permission will receive at most 90% of the grade. Late submission within a week but after the posting of the solution will receive at most 70% of the grade. Other late submissions will depend on instructor's discretion. If you request an extension, you should contact the instructor at least one day before the due day.

Collaboration: You are welcome, even encouraged, to collaborate on the homework and lab assignments, though we urge you to first attempt working out all of the problems by yourself. However, you are expected to write answers **yourself** and understand everything that you hand in. **Copy** results from any sources will be considered as violating Academic Integrity. Collaboration is not allowed on the quizzes and exams.

Classroom disruptions policy: Each Northeastern student has the responsibility to respect the other students in classroom, including using electronic devices. (For example, messaging, game playing, watching videos, and internet surfing are not allowed during class time. However, you can use devices taking notes or quick search some class related information.)

Academic Integrity Policy: Cheating will not be tolerated. All incidents of cheating will be reported. From the Academic Integrity Policy: (see http://www.northeastern.edu/osccr/academic-integrity-policy/)

> "A commitment to the principles of academic integrity is essential to the mission of Northeastern University. The promotion of independent and original scholarship ensures that students derive the most from their educational experience and their pursuit of knowledge. Academic dishonesty violates the most fundamental values of an intellectual community and undermines the achievements of the entire University.
>
> As members of the academic community, students must become familiar with their rights and responsibilities. In each course, they are responsible for knowing the requirements and restrictions regarding research and writing, examinations of whatever kind, collaborative work, the use of study aids, the appropriateness of assistance, and other issues."

Policy about using AI: See Canvas.

Title IX Policy: The University strictly prohibits sex or gender discrimination in all university programs and activities. Information on how to report an incident of such discrimination (which includes sexual

harassment and sexual assault) is located at http://www.northeastern.edu/titleix

Inclusion and Diversity: I value all students regardless of their background, country of origin, race, religion, gender, sexual orientation, ethnicity, or disability status, and am committed to providing a climate of excellence and inclusiveness within all aspects of the course. If there are aspects of your culture or identity that you would like to share with me as they relate to your success in this class, I would be happy to meet to discuss. Also, if you have any concerns in this area or are facing any special issues or challenges, I encourage you to discuss the matter with me as you feel comfortable, with assurance of full confidentiality (only exception being mandatory reporting of NU Academic Integrity Policy violations and Title IX sex and gender discrimination).

Students with disabilities: Students who have disabilities who wish to receive academic services and accommodations should follow the standard Disabilities Resource Center (DRC) procedures (see http://www.northeastern.edu/drc/getting-started-with-the-drc/). College of Science Policies: The current College of Science Academic Course Policies is available at https://cos.northeastern.edu/wp-content/uploads/2012/10/COS-teaching-policies-April-2017.pdf

TRACE: Every student is expected to complete the online TRACE survey at the end of the semester.

Tentative Schedule

Week 1(One day) (September 1-5)

Sec0 Introduction

Week 2 (September 8-12)

Sec1 Probability Theory

Sec2 Conditional Probability

Week 3 September 15-19

Sec3 Transformation

Sec4 Expectation

Week 4 September 22-26

Sec5 Distribution Families

Sec6 Multiple RV

Week 5 (9/29-10/3)

Sec7 Inequalities

Sec8 Random Samples

Week 6 (Oct.6-10)

Sec9 Convergence and CLT

Week 7 (Oct.13-17)

Sec10 Monte Carlo simulation

Midterm Exam 1 (in class)

Week 8 (Oct.20-24)

Sec11 Sufficient Statistics

Sec12 Point Estimation 1

Week 9 (Oct.27-31)

Sec13 Point Estimation 2

Week 10 (Nov3-7)

Sec14 Hypothesis Test 1

Sec15 Hypothesis Test 2

Week 11(one day) Veterans Day

Sec16 Interval Estimation 1

Week 12 (Nov17-21)

Sec17 Interval Estimation 2

Midterm Exam 2 (in class)

Week 13(one day) Thanksgiving Day (Nov24-28)

Sec18 ANOVA 1

Week 14 (Dec.1-5)

Sec18 ANOVA 2

Sec19 Bayesian1

Week 15 (Dec.8-12)

Sec20 Bayesian2

Sec21 Review and Summary

Final Exam