

Statistical Inference - Syllabus Spring 2023

STAT GU4204

March 21, 2023

The aim of this course is to provide a solid foundation in statistical inference. The topics include maximum likelihood estimation, Bayesian inference, confidence intervals, hypothesis testing, linear regression, nonparametric tests, among others. Practical aspects of data analysis, however, will not be covered. Along the way, students will also obtain foundations sufficient for STAT GR5703 (statistical inference & modeling), STAT GU4205/GR5205 (linear regression), STAT GU4291/GR5291 (advanced data analysis), and other more advanced courses.

Note: This syllabus is subject to change, and the most recent version can always be found on CourseWorks.

- **Lectures:** Tuesdays & Thursdays 2:40pm-3:55pm, 702 Hamilton Hall
- **Instructor:** [Johannes Wiesel](#)
- **TA:** Wribhu Banik
- **Instructor office hours & location:** **NEW: Friday, 1.30-2.30pm**, Office 1016, School of Social Work
- **TA office hours & location:** **Friday 11am-12pm**

1 Textbooks

- *Probability and Statistics*, Morris H. DeGroot and Mark J. Schervish, 4th edition
 - Chapters 7–12.
 - **Note:** Suggested problems will be drawn from this text, so the correct edition is required.
- Lecture notes (to be posted on CourseWorks, constantly updated; please let me know if you spot any typos!).

2 Prerequisites

This course assumes a working knowledge of

- calculus (single & multivariable, differentiation, integration, infinite sums, Taylor expansions, limits),
- linear algebra (vectors, matrices, eigenvectors, quadratic forms),
- probability theory (STAT GU4203/GR5203 or equivalent, that is, the content of chapters 1 to 6 of the textbook).

To enjoy the course, you should have a solid interest in the theory and proofs behind the topics discussed in class.

3 Grading

- There will be a homework (HW) problem set due roughly every two weeks.
- The midterm exam will be on **March 09, 2:40pm-3:55pm, during class**.
- The final exam is currently projected for **May 11, 1:10–4:00pm**.
- All students must take the midterm at the time specified above and take the final exam at the time scheduled by the University (projected date/time: see above). **If you have a conflict with any of the exams (for example, due to a religious holiday), please contact the instructor as soon as possible.**
- Exam problems will be similar to those given in the problem sets and worked out in the lectures. You can bring 1 double-sided standard letter size sheet and 1 non-graphing calculator to the exams.
- Class participation and discussion are highly encouraged but not mandatory. The following weights will be used in assigning the final grade:
 - Homework assignments (20%).
 - Midterm exam (40%).
 - Final exam (40%).
- You are encouraged to discuss homework assignments with fellow students, but you must write your solutions **individually** and **by hand**. We accept handwritten solutions on a tablet but **no typed** submissions. Important: for both the midterm and the final exam, you have to write by hand with a physical pen on physical paper. No tablets, computers, etc. Note that:

- Homework assignments will be posted on Thursdays and are due on Friday, 11:59pm, in the following week. Solutions to the assignments will be posted on CourseWorks.
- Homework assignments have to be uploaded to CourseWorks as a **single pdf file**. Phone apps such as Adobe Scan work well to produce pdf scans.
- Note that there are two different dates/times: submit your homework by the **due date** (the earlier one) in order to avoid a late penalty; if you submit after the due date but before the “**available until**” date, your homework will still be graded and counted but a 10% penalty applies. If no homework assignment has been uploaded by the “available until” date, you will receive 0 points for the week.
- Note: CourseWorks applies these rules automatically; the TAs and I cannot and will not interfere. Also, Courseworks is very strict in applying these rules: if the due time is 23:59pm and you submit 23:59:01, your submission may already be flagged as late.
- Do not e-mail me or the TA regarding late homework (we won’t make any exceptions) unless it is due to serious illness or another emergency of similar gravity. In this case, submit proofs (e.g., a doctor’s note) along with your request. Internet issues (unless Columbia or NYC wide), oversleeping, foreseeable absences/appointments or traveling do not count.
- However, your lowest homework score will be dropped before calculating your total homework score for the final grade. It is expected that this will take care of any homework you may fail to hand in on time because of, for example, internet problems, traveling, mild illness, etc. Further accommodations can only be provided because of serious illness or another emergency of similar gravity; a note from a doctor or from a Dean will be required.
- Please verify that you have actually submitted the solutions you wanted to submit (and not a corrupt file, an empty file, another assignment, ...). Submissions can be updated on CourseWorks before the respective deadlines.
- We will not grade all homework questions, but only a subset of the problems, so be sure to check all of your work against the posted solutions. (In short, a 20/20 does not necessarily mean that everything is right.)
- Preparation of solutions (for homework and exams):
 - Unless otherwise stated, you always have to justify your solutions.
 - We may deduct points even if your final result is correct if you do not provide sufficient justification and/or intermediate steps. If you have to compute a quantity, do not simply state the final result. Instead, write down the formula you use, plug in the numbers, and then state the final result.

4 CourseWorks online course system

Class announcements will be sent from and important material is available on CourseWorks:

- You are expected to check the CourseWorks course page regularly.
- A copy of the most recently updated syllabus will be on CourseWorks.
- Lecture notes, assignments and solutions will be posted on CourseWorks.
- Occasionally, there will be other course related handouts posted in CourseWorks.
- You are responsible for making sure that CourseWorks announcements are going to an e-mail you check at least once daily.

5 Campuswire

Campuswire is a Q&A platform designed to get you great answers fast:

- Ask questions!
 - The best way to get answers is to ask questions!
 - Ask questions on Campuswire instead of e-mailing your teaching staff so everyone can benefit from the response (and so you can get answers from classmates who are up as late as you are).
- Edit questions and answers wiki-style.
 - Think of Campuswire as a Q&A wiki for your class.
 - By selecting a category when posting a question, you can specifically ask questions about, for example, “Homework 1”, “Midterm”, or “General.”
- Add a follow-up to comment on or ask further questions.
 - To comment on or ask further questions about a post, start a follow-up discussion.
 - Mark it resolved when the issue has been addressed, and add any relevant information back into the Q&A above.
- Contacting the instructor and the TA:
 - Please use Campuswire for all academic or organizational questions.
 - If you have a *personal matter* to discuss, please go to the instructor’s or the TA’s office hours and request a private zoom session *or* send an e-mail to the TAs. They will handle your request or, if necessary, forward it to me.

To register, please go to <https://campuswire.com/p/GB4B84815> and use the code: 7076

6 Students with disability-related accommodations

In order to receive disability-related academic accommodations for this course, students must first be registered with their school Disability Services (DS) office. Detailed information is available online for both the [Columbia](#) and [Barnard](#) registration processes. Refer to the appropriate website for information regarding deadlines, disability documentation requirements, and [drop-in hours/intake session](#) (Barnard).

Students registered with the Columbia DS office can refer to the Master TARF section of the [DS Testing Accommodations](#) page for more information regarding disability-related academic accommodations for this course.

7 Academic integrity

Columbia's intellectual community relies on academic integrity and responsibility as the cornerstone of its work. Students are expected to exhibit the highest level of personal and academic honesty as they engage in scholarly discourse and research. In practical terms, you must be responsible for the full and accurate attribution of the ideas of others in all of your research papers and projects; you must be honest when taking your examinations; you must always submit your own work and not that of another student, scholar, or internet source. Failure to observe these rules of conduct will have serious academic consequences, up to and including dismissal from the university. If a faculty member suspects a breach of academic honesty, appropriate investigative and disciplinary action will be taken following Deans Discipline procedures. You are encouraged to check the Columbia University Undergraduate Guide to Academic Integrity at <https://www.college.columbia.edu/academics/academicintegrity>.

8 Topics

This is a **tentative** schedule for the course. The exact topics will depend on progress made in class.

- Introduction: Motivation, recap of probability, limit theorems and delta method
- Basics of Statistical Inference: Statistical model, statistics and estimators, method of moments estimators

- Maximum Likelihood Estimators (MLEs): Definition, properties and computational methods
- Principles of estimation: Bias, variance, MSE of an estimator, MVUE, sufficient statistics
- The Bayesian paradigm: Prior and posterior distributions, Bayes estimators
- Sampling distributions: Gamma, χ^2 -, and t -distributions, sampling from a normal distribution
- Theoretical Statistics: Fisher information, Cramér–Rao bound, large-sample properties of MLEs
- Principles of Hypothesis Testing: Critical regions, test statistics, power function and types of errors, significance level, p-value, simple hypotheses, uniformly most powerful (UMP) tests
- Specific Hypothesis Tests: t -test, F -test, likelihood ratio tests, equivalence of confidence intervals and tests
- Linear Regression and Linear Models: Least squares, simple linear regression, multiple regression, ANOVA
- Nonparametric Statistics: Goodness-of-fit tests, categorical data analysis, χ^2 -test, sample distribution function, Kolmogorov–Smirnov tests