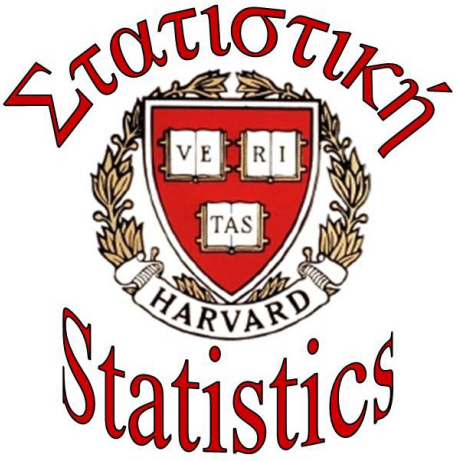
Guidelines for Honors Designations and Senior Theses
Department of Statistics, Harvard University



Academic year: 2026 - 2027

Joe Blitzstein, Director of Undergraduate Studies
Kevin Rader, Associate Director of Undergraduate Studies
Alex Young, Undergraduate Advisor and Lecturer

1 Honors eligibility

For eligibility for an Honors designation, there are two possible paths under the Statistics Concentration Requirements: one based on writing a senior thesis and the other based on following an advanced course option. The possible outcomes for an honors designation in Statistics are no honors, honors, high honors, or highest honors. These are departmental honors (also known as "English honors"), not to be confused with Latin honors (cum laude, magna cum laude, summa cum laude). For details on Latin honors, see the Harvard College handbook.

1.1 Senior thesis

To be eligible for high honors or highest honors, a thesis is required. The senior thesis in Statistics presents an opportunity for a Statistics concentrator to experience

mentored research. This can be a very rewarding experience, giving the student an opportunity to delve deeply into a topic they find interesting, possibly obtaining new results along the way. Gaining research experience by writing a senior thesis is especially useful for students considering graduate school and/or a career in research. A well-written thesis has a cogent, lucid, and engaging discussion both of the literature and of the student's own results.

The writing process provides valuable experience in statistical communication. As such, the use of generative AI to draft any portion of the thesis is not allowed; however, you may use these tools to help proofread your writing. Using generative AI to assist with preliminary research including the identification of related topics, techniques, results, or data is acceptable but should be done with extreme caution. Be aware that in many cases existing software has misstated facts, provided inaccurate references, and in some cases, completely fabricated research results. You are responsible for the veracity of claims in your thesis so careful review of any material provided by AI is critical. Should you have questions about your use of generative AI, please contact the authors of this document.

Since statistics has so many applications and connections to other fields, the Statistics Department also welcomes joint concentrations (with departments that allow this); for a joint concentration, a senior thesis which interweaves the two fields is required.

1.1.1 Honors decisions

Writing a senior thesis is necessary but not sufficient for a High Honors or Highest Honors designation in Statistics. For thesis writers, a GPA of 3.5 or above in Statistics and related courses plus an excellent thesis are expected in order for the student to receive honors (or high/highest honors). The 3.5 GPA is a guideline, not a sharp cutoff; the difficulty of a student's coursework is also strongly taken into account. Students with below a 3.5 GPA should be aware that it is possible not to get honors even if they write a good thesis.

Each thesis has at least two readers: the advisor(s) and, if needed, additional readers chosen by the Department. Suggestions or requests from the student are welcome. The readers are selected, whenever possible, so that their scholarly expertise aligns with the thesis contents. Decisions on honors are made by a vote of the Statistics faculty, based on the recommendations of the thesis readers and performance in Statistics and related courses.

This remainder of this document contains guidelines and advice for senior theses in Statistics, and outlines the department's policies on Honors designations. The precise rules for senior theses and Honors designations are subject to change and are interpreted by the

Statistics faculty.

2 Content and scope

Typically, a senior thesis in Statistics falls into (at least) one of the following (not entirely mutually exclusive) flavors:

1. **Application with strong statistical component:** A statistical analysis of a complex data set that advances understanding in a field other than statistics. While such a thesis may rely primarily on existing statistical methods, it should provide compelling results of interest to the non-statistics field.
2. **Statistics focused on an application:** A modification of and/or empirical comparison of existing methods with a specific applied problem in mind. While the thesis does not necessarily contain an extensive mathematical study of the method(s), it should support the use of the method(s) for the applied problem via some combination of strong plausibility arguments, simulations, and real data analysis.
3. **Statistical methodology:** The proposal of a novel statistical method that addresses one or more applications of interest. The proposed method should be demonstrated (in theory and/or in simulations) to outperform existing methods in some setting(s) of interest.
4. **Statistical or probability theory:** A rigorous theoretical study on questions of interest in statistics and/or probability, such as lower-bounding the attainable statistical accuracy for a statistical inference problem, characterizing the behavior of a statistical method, or proving a limit theorem for a sequence of random variables. A thesis in this area will often contain both a detailed mathematical analysis as well as supporting simulations.

Other options are also possible, with approval from the advisor and one of the authors of this document (Director of Undergraduate Study: Joe Blitzstein; Associate Director of Undergraduate Studies: Kevin Rader; or Undergraduate Advisor: Alex Young). For example, it is possible to write an excellent senior thesis centered around experimental design (supported by mathematical and/or simulation results, especially if the experiment has not yet been conducted).

A good thesis is clear, readable, and well-motivated. A full bibliography must be included. It should justify the applicability of the methods used rather than, say, mechanically

computing a bunch of p-values without checking plausibility of the assumptions or performing diagnostics. Students should discuss with their thesis advisors and/or the authors of this document about what makes for good statistical writing. The junior tutorial (Stat 98) is also highly recommended for any junior who is contemplating writing a thesis.

There is no minimum or maximum length; a cogently-written 40 page thesis is far better than an 80 page rambling thesis! Very roughly speaking, 40 – 80 pages is typical (this estimate is based on using a 12 point font, 1 inch margins, and double-spacing). But the appropriate length depends greatly on the topic (as well as the number of plots and tables, etc.). *Clarity of thought, depth of insight into the research questions, and quality of writing are what make a thesis impressive, not the word count.*

Statistics senior theses have been exploring a wide and ever-growing range of applications, theory, and methods. Some recent topics are listed below, in no particular order. A vast array of topics are possible, with varying combinations of data analysis, experimental design, modeling, inference, theoretical results, and simulation results. If you're choosing a topic, make sure to find something *you* find interesting (so that you'll enjoy working on it), with a clearly-defined research question. Finding a topic can be explored iteratively by meeting with possible advisors. This should be done *early*, ideally during junior Spring, but it common to arrange an advisor during the first few weeks of Senior Fall.

Don't be shy about talking to the authors of this document, professors from your courses, other professors, either in the Statistics Department or outside of it whose work involves a lot of statistics, your TFs, and your fellow concentrators!

2.1 Some recent senior thesis topics

1. When Luck Becomes the Arbiter: Responsibility, Risk, and the Limits of Outcome-Based Judgment
2. Quantifying the Operational $PM_{2.5}$ Burden of Hyperscale Datacenters
3. Accelerating Inference: Mitotic Stein Variational Gradient Descent for Bayesian Analysis of Dynamical Systems
4. Paint and Pixel: Reimagining Impressionist Color in Monet's Rouen Cathedral Series through Statistics
5. Epistasis between Somatic Mutations in the Normal Blood
6. Bootstrap Methods for High-Dimensional Linear and Logistic Regression
7. A Multimodal Foundation Model for Contextualizing Protein Function

3 Advisors

Each senior writing a senior thesis in Statistics should have an advisor or co-advisor who teaches in the Statistics Department or Biostatistics Department. In some cases, students writing a thesis with substantial content from another field (e.g., CS, Economics, Government, or MCB) can have a faculty member from this field as primary advisor or co-advisor; such an advisor must be approved by one of the authors of this document. A student requiring help with finding an advisor should meet with the undergraduate advisor leading the Stat 99r tutorial (Alex Young).

Students pursuing a joint concentration in Statistics are expected to have an advisor or co-advisor in the Statistics Department. Exceptions can be made with approval from the authors of this document. For questions or guidance on this issue, please contact . Note, that the second concentration may have its own requirements for the advisor(s). Ordinarily and preferably, the advisors are chosen by the student.

4 Timeline

The most successful thesis experience generally follows a disciplined schedule that provides plenty of time for research, reflection, and writing, while giving the advisors sufficient time to comment on drafts and the student sufficient time to make improvements based on the comments. Keep in mind Hofstadter's Law: "It always takes longer than you expect, even when you take into account Hofstadter's Law."

A suggested timeline for thesis-writing is below. There are only two required submissions (in **bold** on the timeline): the October 1 proposal and the March 26 final thesis submission. The rest of the dates on the timeline below are suggestions only; you should work with your thesis advisor(s) to come up with an individual schedule with regular milestones such as completing drafts of certain chapters.

- Junior year: student chooses a possible thesis topic and advisor or advisors. The junior tutorial in Statistics (Stat 98), may help with this process.
- Summer between junior and senior years: thesis research often begins, though it is also feasible not to start work on the thesis until senior fall.
- **October 1**, senior year: thesis proposal submission by 11:59PM. The proposal is a brief, one to two page outline of the main questions to be studied, the data or other resources that will be used, and the student's outline of the timeline for writing the

thesis. Be sure to also specify on your proposal who your thesis advisor(s) will be. In particular, this means you need to have an advisor lined up well in advance of October 1st (preferably by the end of junior spring, or at least by the start of the fall semester). Please submit your proposal through the Stat 99r Canvas page. Contact the 99r instructor (Alex Young) if you need guest access to the course page.

- November 10: literature review submitted to advisors. The literature review will consist of a brief outline of relevant literature in the field of the thesis. Typically, this review serves as the first chapter in the thesis, perhaps with some later revisions, and will be approximately 15 – 20 pages in length (including the list of references). While the review should contain an extensive bibliography, it is much more than a list of published papers. A useful review clearly summarizes the essential results of the most relevant papers in a coherent, organized manner and discusses how these papers will influence or inform the research.

- January 5: detailed outline of thesis submitted to advisors. If a thesis will involve substantial data analysis, this outline should be accompanied by initial analyses, which may include simple data tables, graphics or more sophisticated models. If the thesis will involve simulation, the outline should include an outline of the design of the simulations and some early results from the simulations.

- March 1: complete draft of the thesis due. The draft may be in rough form, indicating where material will be added, but should contain approximately 80% of the material that will eventually be in the thesis.

- **March 26**: completed thesis submitted electronically to the Stat 99R Canvas page by 5:00 pm, in PDF format. If you are not enrolled in Stat 99R, please email the Stat 99R instructor well in advance (preferably at the start of the semester) and ask to be added as a "guest" on the Canvas site, which will give you full access to the materials posted on Canvas, as well as access to the submission page.

4.1 Off-cycle deadlines

For off-cycle concentrators (i.e., students who will graduate in a fall semester rather than a spring semester), the proposal is due by March 1 if the student's second-to-last semester is a spring semester or October 1 if the student's second-to-last semester is a fall semester. The off-cycle senior thesis submission deadline is November 6th at 5:00 pm.

5 Format

The thesis should be neatly typed, with a font size of 11 pt or 12 pt. For the typesetting, using LaTeX or R Markdown is highly recommended but not required. A LaTeXtemplate is available on the Stat 99r Canvas page. Joint concentrators should check with their other concentration for specific formatting requirements.

References that were used should be cited, in the main text and in a bibliography. Of course, the thesis should comply with the Harvard College Honor Code and the policies on academic integrity described within the Harvard College handbook

The thesis should also include an acknowledgements section (to mention people and institutions that provided various forms of support) and an abstract briefly summarizing the main motivation, problem, and results.

Electronic submission is required by 5:00 pm on March 27th through the Stat 99R Canvas website. Please submit your thesis as a single PDF file. If you wish to include code include a link to a shared repository (e.g. github) containing the files. **This is a hard deadline; no extensions are granted on thesis submission.**

Hard copies are not required, but many advisors appreciate receiving a hard copy (as well as being a nice addition to your own bookshelf). Hard copies should be printed on acid free paper and securely bound. There are various nice ways to do the binding; for example, the binding can be done at Gnomon's or with the black spring binders sold at the Coop.

6 Stat 99R

As a Statistics senior thesis writer you can, but are not required to, register for Stat 99R for 1 or 2 semesters in which you are working on you thesis. Stat 99R provides course credit for working on a thesis, and meets in person once approximately weekly to check in on everyone's progress, offer support and answer questions on the thesis process, and foster a community of thesis writers.

Stat 99R can be counted *once* for concentration credit, as a *related* course. So if you take Stat 99R twice, it counts once as a related course for the concentration and once generically for Harvard course credit. Stat 99R is graded SAT/UNSAT, but does *not* count as your one pass/fail or SAT/UNSAT concentration course. That is, you can still take one course (other than Stat 110, 111, 139) for concentration credit pass/fail or SAT/UNSAT, in addition to taking Stat 99R.

If you are a joint concentrator and the other side of your concentration has its own senior thesis tutorial, you should only register for one of the two thesis tutorials per semester.

If you are writing a thesis but are not officially enrolled in Stat 99R, you are still welcome to participate in the tutorial informally, and you should ask the Stat 99R instructor to add you as a "guest" on Canvas.

7 Thesis Advice from Statistics Alumni

Some reflections on the thesis-writing process from statistics concentrator alumni are below. There is a consensus that one should start early and make sure to find an exciting topic and a supportive advisor, and that it's a challenging but rewarding experience.

1. *I really enjoyed my thesis writing experience, although at times, it seemed like more of a love-hate relationship. It was tough to force myself to sit down and write, especially when my non-thesis writing friends seemed to be having so much more fun that I was. However, writing a thesis was overall a very rewarding experience. It gave me the opportunity to do real-world statistics research (and since I'm not going into academia, this was really my only chance) and apply the things that I learned in classes that I have taken throughout my time at Harvard.*

After months of hard work on the thesis, having the final product really gave me a sense of accomplishment–like I had done something really worthwhile with my senior year. So while the process of writing the thesis can be a real drag at times, I feel that in retrospect, it was a really great decision for me. My only advice to future thesis writers would be to make sure to start early. Ideally, have a topic before the end of junior year and work on the thesis over the summer a little bit so that your feet are at least a little wet by the time the real work begins during senior year.

2. *For future concentrators, I would suggest working on research with one of the professors much earlier than junior summer, and [participating in research meetings] so that there is plenty of time to decide whether or not to do a thesis. Working on something I thought was worthwhile and novel was much more motivating to my thesis writing than any other factor.*

3. *If you are considering doing research after college, the senior thesis can be a chance to find out whether a particular area of statistics motivates you and to gain a baseline knowledge of the literature in that area.*

I continued on to PhD studies in statistics, and writing my thesis gave me an area of focus before I even entered graduate school. Identifying a general research interest early was an important advantage: the skills I picked up while writing my thesis allowed me to begin consulting and collaborations leading to my dissertation right away. Having a specialty also meant that I was able to assist with teaching more advanced courses and answer questions from other (older!) graduate students. I imagine that the same advantages exist if your thesis focuses on an applied topic and you are considering graduate studies in that field.

If you find out that you'd rather be working in a different specialty, you have learned something important. And, if you learn from writing your thesis that research is not your interest, that is worth discovering before you sign up for graduate school! Expertise in a specific field can be an advantage in any post-college pursuit (and can set you apart at job interviews).

4. *Writing a thesis is probably the best decision I made in my college career, and I can boil down the "why" into four main reasons:*

(i) It's fun! In the academic sense, that is. After so many years of by-the- book, structured education (not that there's anything wrong with that), it's refreshing to learn without problem sets. I was free to explore off on any tangent that piqued my interest, with as little or as much instruction as I needed. It's exciting, to get to build your own "coursework" based solely on what interests you.

(ii) I learned a ton. There's the surface stuff - I got way better at R and Latex, for example. But more importantly, I now have an "area of expertise." It's cool to have a very in-depth knowledge of one small corner of statistics, and to feel like the expert in my (incredibly narrow) field.

(iii) It's a good trial run for graduate school. Schoolwork and research are totally different animals, and you don't know which one you will like until you try both. I was pretty sure, but not 100% convinced, that I wanted to get a Ph.D. - after writing a thesis, and discovering how much I enjoyed research, I now am certain. You don't want to jump into something as big as Ph.D. research unless you're sure it's your cup of tea.

(iv) Personal pride. I can't say I put is much into my classes as I could have, but I'm leaving college with a very real academic achievement: 120 pages, gloriously bound on acid-free paper, of my research. That's a cool feeling, I promise you.

In summary: Thesis writing let me get a jump start on real research, with the resources of Harvard, a top Statistics department, and a fantastic advisor all at my disposal. That's quite a unique opportunity, and I'll always be glad I took advantage of it!