PHIL 174: Philosophy of Statistics

last updated: June 25, 2021

Description. We'll survey some special topics in the philosophy of statistics: simplicity, algorithmic fairness, merely statistical evidence, correlation versus causation, randomized trials, and statistical learning theory. On the one hand, we'll see how general philosophical problems (for example: the problem of induction from epistemology, or of defining causation from metaphysics, or of analyzing discrimination from ethics) show up in statistics, and how you might try to solve them in that setting. On the other hand, we'll draw out and evaluate the philosophical assumptions lying behind statistical practice.

The class will be challenging: we'll move fast, covering a lot of ground, and will need to get to grips with sophisticated concepts, notations, techniques, and arguments. But we'll learn these together from scratch. I will assume bits and pieces of high school math. But if you're rusty on any of them, just let me know and I can help you brush up. Anyone ready and willing to hit the ground running will be able to do well in the class.

Pre-requisites. PHIL 21: Skepticism and Rationality, or permission of instructor.

When and where. Tuesdays and Thursdays, 11am–12.15pm, on Zoom. Do your best to make the live class—it's more fun for everyone that way—but I'll post recordings and my slides on the class site for those who can't make it.

Strategies for success. Some strategies are obvious: ask questions, start the homeworks early, make use of office hours, collaborate, and so on. But other strategies are not obvious. Learning—whatever the subject matter—is itself a skill, and a difficult skill, but a skill we can get better at. Try to reflect on how you learn, experiment with different strategies, and discuss your experiments with others.

One strategy I recommend is spaced repetition using an app such as Anki. Spaced repetition is "one of the most powerful, reliable, and easy to use" techniques in learning (Benedict Carey, *How We Learn*). I use Anki and find it invaluable. I encourage you to try it out!

Instructor. Cosmo Grant, cdg@ucla.edu, office hours Mondays 2–3pm or by appointment on Zoom.

Materials. Everything will be posted on the course site.

Readings. Most classes have a reading, which you should look over beforehand. The readings are short but often tricky. Do read them but don't worry if you're confused: the aim is to prime you for class, where we will go over the concepts more slowly together. View them as a challenge: even if the details are confusing, can you extract any key ideas? That's an invaluable skill, and takes practice.

Assessment.

Homework, 80%. There will be five homeworks, due on Gradescope. See below for the schedule. They will be a mix of questions: some testing recall or reading comprehension, some requiring a calculation or argument, some requiring a few paragraphs of prose. Your lowest-scoring homework will be dropped and the remaining four weighted equally. Collaboration is encouraged, as long as you list your collaborators and write up your answers yourself.

Forum posts, 20%. By the end of each week, you should post on the class forum. Your post can be about anything related to the course: a question about the readings, a useful resource you found, something you're confused about, a reply to another post, or so on. They can be short (a few sentences) but should be substantive. The aim is to encourage you to engage independently with each other and the material: philosophy is a conversation, not a monologue. Graded for completion only.

Policies.

Late work. Everyone gets four free extension days for the homeworks, e.g. 11am Thursday becomes 11am Friday. You can use these whenever you like and don't need to give any explanation. I'll apply your extension days to any late homework automatically. If you need an extension beyond these, get in touch and we can talk. If you don't have an extension then homework will be marked down 10 percentage points for each day late (e.g. homework due 11am Thursday but received 12pm Friday will be marked down 20 percentage points). Extension days only apply to homeworks, not forum posts. The aim is to strike a balance between being flexible and helping people stay on track.

Academic integrity. There is zero tolerance for plagiarism and other forms of academic misconduct. In short: cite your sources and list your collaborators. If in doubt, just ask. Asking is the easiest way to avoid any problems.

Accommodations. If you need any accommodations, just let me know.

Schedule.

T March 30. Preview **H1 out**

no reading

R April 1. Simplicity: the problem described

Baker, *SEP: Simplicity*, Section 1

T 6 April. Simplicity: a Bayesian approach

White, *Why Favour Simplicity?*

R 8 April. Simplicity: an alternative approach **H2 out**

Forster and Sober, *How To Tell When Simpler, More Unified, Or Less Ad Hoc Theories Will Provide More Accurate Predictions*, up to Section 3

T 13 April. Simplicity: finishing up **H1 due**

no reading

R 15 April. Algorithmic fairness: the set-up

ProPublica, *Machine Bias*

T 20 April. Algorithmic fairness: the impossibility results

no reading

R 22 April. Algorithmic fairness: a closer look **H3 out**

Hedden, *How (Not) to Test for Algorithmic Bias*

T 27 April. Merely statistical evidence: the problem described **H2 due**

Thomson, *Liability and Individualized Evidence*, up to Section IV

R 29 April. Merely statistical evidence: the cutting edge (guest lecture by Haley Schilling)

reading tbd

T 4 May. Correlation over causation?

Mayer-Schönberger and Cukier, *Big Data*, Chapter 4

R 6 May. Causation defended **H3 due, H4 out**

Pearl, *The Art and Science of Cause and Effect*

T 11 May. Causation explained?

no reading

R 13 May. Randomized trials: arguments for

Howson and Urbach, *Scientific Reasoning*, Sections 6.a and 6.b

T 18 May. Randomized trials: evaluating the arguments

Howson and Urbach, *Scientific Reasoning*, Section 6.d

R 20 May. Randomized trials: ethical issues **H4 due, H5 out**

Royall, *Ethics and Statistics in Randomized Clinical Trials*

T 25 May. Statistical learning theory: the set-up

Kulkarni and Harman, *An Elementary Introduction to Statistical Learning Theory*, Chapter 1

R 27 May. Statistical learning theory: learning from examples

Kulkarni and Harman, *An Elementary Introduction to Statistical Learning Theory*, Chapter 6

T 1 June. Statistical learning theory: universal consistency

Kulkarni and Harman, *An Elementary Introduction to Statistical Learning Theory*, Chapter 7

R 3 June. Review **H5 due**

no reading