Teaching Dossier

Michael Schweinberger

1 Overview of teaching

1.1 Courses developed

I have developed a graduate-level course that introduces students to three popular streams of research in statistics, machine learning, and artificial intelligence:

- networks representing data structure: e.g., data on connections among users of Facebook, Twitter, and LinkedIn;
- networks representing model structure: e.g., models of the human brain and gene interaction networks arising in the study of genetic diseases;
- networks representing mathematical and computational operations: e.g., neural networks (artificial intelligence).

I have taught the course 8 times since 2014.

1.2 Courses taught

I have taught 20 statistics classes at the following universities:

- Department of Statistics, Penn State University (PSU): 2 classes;
- Department of Statistics, Rice University (Rice): 16 classes;
- Department of Statistics, University of Missouri (MU): 2 classes.

These 20 classes include introductions to the mathematical foundations of statistical learning from data at all levels:

- Bachelor's level (8 classes): Stat 401 at PSU (2 times); Stat 310 & Econ 307 at Rice (5 times); and Stat 4710 & 7710 at MU (1 time);
- advanced Bachelor's and Master's level (3 classes): Stat 419 & 519 at Rice (3 times);

- Ph.D. level (9 classes): Stat 532 at Rice (1 time); Stat 648 at Rice (7 times); and Stat 9100 at MU (1 time).

1.3 Ph.D. qualifying exam

As a result of teaching the core courses on the mathematical foundations of statistical learning from data at Rice (Stat 519), I was in charge of preparing and grading the written Ph.D. qualifying exam in statistics at Rice and leading oral examinations in 2018, 2019, and 2020.

2 Teaching philosophy

To prosper in an uncertain and changing environment, humans and machines (artificial intelligence) need to learn from observations. Statistics is the science of learning from observations and can therefore serve science and human welfare by studying from first principles how humans and machines can best learn from observations in an uncertain and changing environment, and by providing mathematical and statistical guarantees for statistical learning procedures.

As a consequence, statistics has much to offer to students, but the age of data science presents both opportunities and challenges for teaching statistics. On the one hand, it presents opportunities in that statistics courses have witnessed a surge of demand and statisticians have the opportunity to help shape an important slice of the world's future workforce: data scientists. On the other hand, it presents challenges as statistics courses have more students, and the background and career goals of those students is more heterogeneous than in the past: e.g., many undergraduate students enrolled in statistics courses do not major in statistics, but major in computer science, engineering, the natural sciences, or the social sciences.

I have attempted to make my courses useful for all students, regardless of background and career goals, and have placed a strong emphasis on

- connecting with students by choosing real-world examples that appeal to them, such as examples from the health sciences, machine learning, artificial intelligence, and other fields;

- leaving no student behind, making sure that all students understand the key ideas of how to learn from data.

2.1 Connecting with students

In all of my courses, I attempt to connect with students by choosing real-world examples that appeal to them and by inviting students to ask me questions before, during, and after class.

Many of my real-world examples are taken from three popular streams of research in statistics, machine learning, artificial intelligence, and related fields:

- networks representing data structure: e.g., data on connections among users of Facebook, Twitter, and LinkedIn;

- networks representing model structure: e.g., models of the human brain and gene interaction networks arising in the study of genetic diseases;

- networks representing mathematical and computational operations: e.g., neural networks (artificial intelligence).

Examples taken from the health sciences, machine learning, artificial intelligence, and other fields highlight the importance of statistics in the age of data science, because machines and neural networks need to learn from observations in order to prosper in an uncertain and changing environment, and statistics is the science of learning from observations. I therefore select simple examples from the health sciences, machine learning, artificial intelligence, and other fields whenever appropriate. As a case in point, when introducing probabilistic models of two variables, I introduce a simple probabilistic model of a human brain with two neurons that are either activated or not activated while performing a task (e.g., walking). I explain the probabilistic properties of the model step by step, support my explanations by numerical examples, and demonstrate that the model can provide interesting insights into how the human brain operates. A related example is a simple probabilistic model of an artificial brain with two artificial neurons, which serves as a simple example of a neural network in artificial intelligence. I choose these examples with care and make sure that the examples are as simple as possible and do not distract from the main contents of the course, but make students curious about statistics, the science of learning from observations.

2.2 Leaving no student behind

I have designed my courses so that no student is left behind, regardless of background and career goals. To leave no student behind, I adhere to the following principles:

- At the beginning of my classes, I make sure that all students are on the same page, by assessing the knowledge of students and providing remedial resources, as needed.

- I place a strong emphasis on the key ideas of how to learn from observations. I believe that the key ideas are simple and that every student can understand them. I make the key ideas as accessible as possible and support them by careful mathematical arguments that start simple and proceed step by step, without skipping steps that may seem trivial to me but may not be obvious to students. I encourage students to ask questions during class, so that I can address gaps in understanding on the spot. Since not all students are comfortable asking questions during class, I invite students to ask me questions after class.

- I take advantage of graphical representations whenever possible, in order to help students with a weak mathematical background understand the key ideas and make them more comfortable.

- In collaboration with Minjeong Jeon (Graduate School of Education & Information Studies, UCLA), I am developing statistical learning progression maps for monitoring the learning progress of students and helping detect students who need more support than others. We have developed a prototype of a statistical software for providing statistical learning progression maps and I intend to use statistical learning progression maps in my future classes to

  — assess how much progress students make;

  — assess whether and how much students can improve;

  — detect students who need more support than others.

  An example of a statistical learning progression map — based on data collected on the online learning platform My Math Academy — can be seen in Figure 1 on the following page.

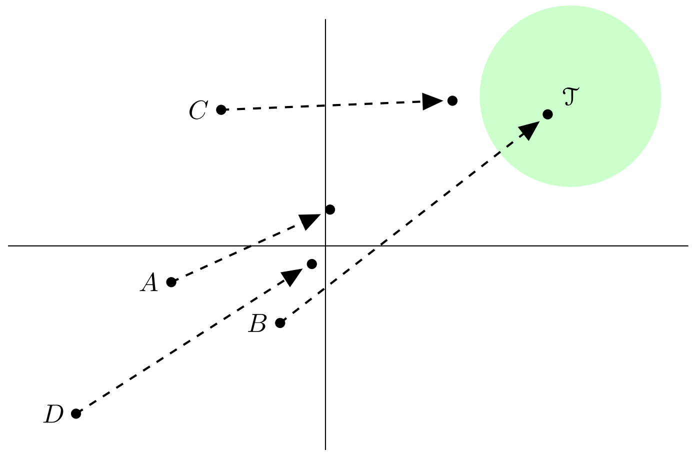

Figure 1. A statistical learning progression map based on data collected on the online learning platform My Math Academy. Students A, B, C, D, and other students took an online test with 30 problems at two time points. The test measures progress towards learning target $\mathcal{T}$, the understanding of numbers and arithmetic operations. The statistical learning progression map reveals that some students made more progress than others, and some of them reached the target while others can still improve.

2.3 Undergraduate courses

In undergraduate courses, I have found it useful to take advantage of analogies, such as analogies with criminal investigations and criminal trials. For example, to introduce statistical estimators, I use analogies with criminal investigations: Statistical estimators resemble detectives in criminal investigations, using a trail of evidence to track down the source of the evidence. To explain statistical tests, I use analogies with criminal trials.

3 Teaching evaluations

I have taught 16 out of 20 classes at Rice University, where I was a tenure-track Assistant Professor from 2013 until 2021. I therefore present all teaching evaluations from Rice University from 2017 until 2021, which are 24 pages. All of them are student-based teaching evaluations, because Rice University does not issue peer-reviewed teaching evaluations.