Teaching Statement

Fan Bu, Ph.D.

> "Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write." — H.G.Wells

Learning Statistics

Learning statistics is the internalization of the principles of statistical thinking and best practices in data analysis. In a classroom, students should be exposed to and also personally go through the entire investigative process of evidence-based problem-solving and decision-making, which is the central goal of statistical analysis.

Thanks to the rising popularity of Data Science in recent years, students from increasingly diverse academic backgrounds and learning interests are now taking classes in Statistics. It is essential as an educator to accommodate this wide range of knowledge bases and learning needs, introduce more hands-on experiences to facilitate long-term retention of practical skills, and communicate in the "students' language" with creativity and clarity.

Ensuring Equity of Learning

Given the diverse array of students in our courses, particularly at the introductory level, it is fundamental that we learn about students' knowledge gaps prior to the beginning of the semester. This is critical in that it offers us the opportunity to understand what students know and the practical steps that can be taken to pedagogically meet them where they are at. For example, in my introductory undergraduate-level course, *Data Analysis & Statistical Inference* [a], I conducted **anonymous** surveys on students' knowledge bases in mathematics and programming skills. Similarly, in an introductory graduate-level course, *Bayesian Methods and Modern Statistics* [b], as a teaching assistant, I administered **ungraded** quizzes to gauge students' mathematical ability. Such informal knowledge checks can inform how much calculus and probability theory I need to cover, and whether or not I need to set aside some lecture time to walk through the basics of statistical software such as `R`, `Python` or `SAS`.

[a] Course website: `https://sites.google.com/view/sta101-001-summer2021/home`.

[b] Course syllabus: `http://www2.stat.duke.edu/courses/Fall19/sta601.001/syllabus_601.pdf`.

Another focus in my pedagogy is finding the right balance in how I teach remedial and advanced students. For students who struggle with the material, I encourage them to attend office hours and to study the available lecture notes and slides, and I am also open to one-on-one email correspondence and meetings to provide extra guidance. In addition, I begin each class with a summary of the important concepts and techniques, as well as devote 10 minutes to answering questions that emerged during homework assignments. To keep advanced students motivated, I provide optional challenges for extra credit. For example, when I taught the *2018 Duke Statistical*

Science Bootcamp [c], I was asked to review the pre-requisite knowledge to prepare new graduate students for their upcoming coursework. Due to the nature of the experience, we start with the basics (such as multivariate calculus and linear algebra), compile a "daily challenges" problem set [d] , and encourage students to attempt the problem set when they have finished the exercises. While not all students complete the problem sets, they do stimulate collaboration and peer learning. Towards the end of the Bootcamp, the problem sets were colloquially known as the "happy hour" of the bootcamp.

[c] The bootcamp learning materials are available at `https://github.com/fanbu1995/DukeStatSciBootcamp2018`.

[d] The challenge problems are compiled at `https://github.com/fanbu1995/DukeStatSciBootcamp2018/blob/master/Challenge_Problems.pdf`.

Combining Hands-on Sessions with Theory

Statistics isn't all about probability theory and mathematical derivations. No matter how elegant the Bayes rule or the central limit theorem is made out to be in lectures, students still tend to ask, "so what?" Theory only paves the road of learning; it is during practical problem-solving experiences that the students truly take on the journey. In my undergraduate-level introductory course, *Data Analysis & Statistical Inference*, I designed hands-on lab sessions with materials relevant to concepts covered in class. For example, during the basic probability theory week, students used real NBA basketball data to investigate the "hot hand" phenomenon of top players like Kobe Bryant and tested if future shots made are dependent on previously made shots by computing conditional probabilities [e]. Such carefully selected case studies on entertaining topics often resonate well with students and help them grasp otherwise elusive concepts. [f]

[e] GitHub repository for the lab materials: `https://github.com/fanbu1995/sta101-materials/tree/main/Labs/lab3`.

[f] In course evaluations, a student reflected upon their experience with the lab sessions, "*it allows you to know the boundaries and guidelines of the things that you need to pay attention to when conducting your own test for your own projects, [and also] the correct interpretation of what you get after computing [a statistic]. It's not just math. It's poetry too.*"

Communicating with Creativity and Clarity

Exposition strategy is central to effective teaching. Certain statistical concepts such as "p-values" and "confidence intervals" are challenging to understand, so when repeating the textbook definition doesn't work, using metaphors, drawing diagrams or running simulations might. That being said, getting across the message requires some trial-and-error; one exposition technique may work well for one student, but not another. From my experience, a professor should strive to take a more personalized approach and try to "speak the student's language". I have been honing my skills in this aspect for years as a private Statistics tutor in addition to office hour duties, and I'm constantly learning new ways to elicit that precious "oh!" reaction.

On the other hand, communication with clarity is crucial for establishing expectations, which directly influence the quality of experience and how much (or little) the students learn and retain. It is critical that a professor provides clear and consistent expectations about their classroom. In my experience, students want regular feedback on their assignments and progress, clarity in terms of what is being asked of them, and transparency with respect to how they will be assessed. A clear set of standards, posted upfront, alleviates most issues before they arise. In *Data Analysis & Statistical Inference*, for example, I provided a detailed grading rubric [g] for the final course

[g] Details on the course project page: `https://sites.google.com/view/sta101-001-summer2021/project`.

project so that the students would know what was expected and how to structure their presentation and report. Furthermore, homework solutions [h] were uploaded to the course website to help the students reflect on their work. In *Statistical Case Studies*, a graduate-level course on data analysis and scientific communications, we provided a guideline for writing the case study reports [i] and offered editorial comments when grading reports, so that students could assimilate the standards of academic writing.

[h] A compilation of homework solutions is available here: `https://github.com/fanbu1995/sta101-materials/tree/main/Problem-sets-key`.

[i] The written report guideline can be seen here: `https://fanbu1995.github.io/teaching/Case_Studies_Written_Report_Guideline.pdf`.

Self-development and Mentoring

Given my pedagogical experiences in diverse student bodies and course designs with hands-on sessions at both undergraduate and graduate levels, I am prepared to teach a variety of statistical courses. In particular, I am very interested in teaching introductory-level courses aimed at first- or second-year statistics majors or non-majors who wish to learn statistical methodologies while gaining practical analytical skills. In order to further develop my pedagogical skills and obtain more training in higher education, I have completed the *Certificate in College Teaching* [j] training program at Duke University.

[j] Program webpage: `https://gradschool.duke.edu/professional-development/programs/certificate-college-teaching`.

Furthermore, I have been an active mentor in the past years. During the past summer, I mentored an undergraduate student in a quantitative science research program at UCLA and guided her through an empirical study of Bayesian sequential testing methods for vaccine safety surveillance; [k] it was a very enjoyable and fruitful experience for me and my student, which earned my student a Research Excellence award and me an Excellent Mentor award. Before that, I had mentored several undergraduate students in various collaborative research projects [l] , and also served as a senior student mentor for first-year PhD students to help them transition to a PhD program. In my experience, mentoring is about helping the mentee find their own interests and paths, and providing them with resources and tools for their own intellectual growth. I have found mentoring tremendously rewarding, as working with junior researchers has made me a better communicator and educator, and also inspired me to reflect upon and expand my own research. I have developed a great passion for mentoring, and I am confident that I can apply and extend my past experience to succeed in future mentoring roles.

[k] Results of this study are presented in an RShiny App at `https://kariexue.shinyapps.io/BayesianSimulationResults/`.

[l] For example, I served as the project manager and mentor for *Duke Data+ 2019*, and as a consultant and mentor for *Duke DataFest* from 2018 to 2020.