Teaching Statement

Fangyi Wang
Department of Statistics, The Ohio State University

My teaching philosophy begins with a simple question: *What does this course look like from a student's perspective?* I aim to understand students' backgrounds and goals, clarify the specific takeaways I want for them, and design instruction that helps them reach those outcomes.

I learned this perspective first as a student. In the first semester of my master's program, I took regression while coming from a biology background. Two weeks in, we were discussing the sampling distribution of the OLS estimator, yet I had not formally seen the definition of expectation, which came later in a probability course I was taking in parallel. That experience reminds me how easy it is, after years of learning, researching and applying statistics, to forget the challenges of a first statistics course. Hence, as an instructor, I try to temporarily forget my advanced training, and step into the shoes of novice learners.

How to make this concrete? I use two complementary routes: *top-down* (making abstract ideas concrete) and *bottom-up* (connecting practice back to concepts).

I. Top-down: making abstract math understandable

Statistics can feel opaque with rich notation and few intuitions, especially to students from non-statistics backgrounds. Therefore, when introducing new concepts and notation, I pair them with specific examples and make explicit translations from symbols to numbers in real-world applications. Early in my time teaching Introduction to the Practice of Statistics (STAT 1450) to non-majors, I present the sample mean and variance with both formulas and data examples, such as Big Ten application fees. After the class, a student asked, "What are the x_i?", which calls my attention to not treating any notation as obvious: I need to write explicitly that x_i denotes application fee for each university (each sample). Moving forward, I always pair notation with plain-language labels and a visible mapping between them.

Later in the semester, we start to introduce more theoretical ideas such as population versus sample, weak law of large numbers, sampling distributions and central limit theorem. These topics are abstract but important, since it lays the foundation of constructing confidence interval and testing statistical hypothesis. To make these intimidating theories more friendly and accessible, I use simple and intuitive demonstrations, e.g., coin flip webpage and short simulations in `R` to show empirical convergence and approximate normality. These translate statistical terminologies (math language) to real-life scenarios and pictures (human language).

However, what has been learned in class only starts to click when students understand how to apply them to problem-solving. To encourage this, I resist "spoiling" answers when discussing homework. Instead, I point students to the relevant definitions or procedures in the lecture notes and show how the problem fits into that framework. As a teaching assistant, not knowing an instructor's exact presentation made this harder; as an instructor, I intentionally reference similar examples from slides and textbook so that students can know "how to fish" instead of "getting a fish".

II. Bottom-up: connecting implementation and real problems back to concepts

Students often struggle to recognize which concept applies in a word problem or a coding task. For instance, distinguishing a z-test from a t-test depends on whether the population standard deviation is known. Even though students manage to memorize this difference, it may not always be clear to them when the problem has a lot of words and numbers. Similar confusion shows up in sample mean versus population mean. I notice these confusions in exam grading, where common mistakes are shown repeatedly. Therefore, after each midterm, I'd write a short summary and revisit key concepts in class, e.g., explicitly tying z-test and t-test back to their respective assumptions and conditions. This closes the loop between assessment and instruction.

I had similar experience when teaching recitations for Introduction to Statistical Inference for Data Analytics (STAT 3202). In this class, many students are new to coding. I have seen confusion between sample size and Monte Carlo replicates. To help with these, I briefly review the concepts from textbook before live-coding and try to keep coding variables consistent with the symbols in the notes, e.g., `n` versus `B`. This deepens the connections between keyboard and textbook. In general, connections that are natural to experts may be much less obvious for learners. Making those links explicit is part of the lesson.

Apart from building connections between theory and practice, telling real-world stories itself is enjoyable for both students and instructor. Therefore, I always enjoy teaching regression and experimental design, as they are tied with both fruitful scientific questions and daily life scanrios—the example we used for simple linear regression is eating breakfast and college GPA. In that class, a statistician chose to put "association versus causality" aside and savored the moment of discovering positive correlation together with students. Towards the end, I emphasized association *does not* implies causation, which may have broken a few hearts who planed to eat more breakfast.

Courses I plan to teach

With comprehensive teaching and research experience, I am prepared to teach a range of undergraduate courses, including introductory statistics and data analysis, probability, statistical inference, regression and linear models, introductory statistical learning, etc. At the graduate level, I can contribute to courses in nonparametric statistics, statistical learning, functional data analysis, and special topics related to distribution-free methods and conformal prediction. I am happy to tailor contents based on the program while maintaining clear learning objectives and inclusive pedagogy.