Becoming a Statistician

Abbe Herzig

One of the best things about being a statistician is the broad variety of places where statistics is used. As an applied statistician, each new project required me to learn something new: a little taste of medical research, air quality, engineering, social policy, law. Sometimes I've had to dive into statistical theory in order to apply or even develop appropriate methods. This variety has kept me learning and never let me get bored. Unlike some of my friends who dreaded Monday morning, I've (almost) always been eager to start the work week because my work was always changing and challenging, and there were interesting things to do. I have also learned a few things about building a career as a statistician.

Look beyond the coursework. When I was in graduate school, I didn't know quite what statistics was as a subject of study or as a profession. I earned a master's degree in statistics because I wasn't sure what I wanted to do and one of my professors suggested it. The coursework didn't really excite me, and I thought that once I finished my degree I'd find something else to do. After all, any master's degree is a good job qualification, whether or not it is directly related to the job (this is the case in all of the mathematical sciences, though not for all types of jobs). At the end of my last semester, with the end in sight, my adviser invited me to assist him as a consultant to the Legal Defense Fund of the NAACP, analyzing salary data for an employment discrimination lawsuit against a major US firm. Through that project, I was doing work that would directly help people, and the work gave me a glimpse into the workings of a lawsuit, which captured my attention in a way that my coursework didn't.

That is when I began to discover the many interesting things that my mathematical and statistical courses prepared me for. As in many fields, coursework provides necessary training, but it doesn't always reflect the beauty and fascination of the work.

You don't need to know where you'll end up, just where you want to begin. Statistical and mathematical thinking are needed practically everywhere. Some of the largest employers of statisticians are pharmaceutical and insurance firms, but there are also many other employment options, particularly given the ubiquity of "big data" and data science.[15] As you gain experience in statistics, your skills and knowledge can be applied to an increasing number of things, and the path forward can take unexpected turns.

My first job after graduate school was at Brookhaven National Labs, where I provided statistical support for medical research. When I went back to graduate school for a PhD, one of my professors was the lead statistician in the consumer product testing division of *Consumer Reports* (CR). At the end of the semester I got the courage to ask him for some consulting work, and that led to many years on the staff of CR, analyzing data from large-scale surveys on the reliability of cars, and designing, analyzing, and interpreting product tests ranging from car batteries to pantyhose to lawn mowers. CR eventually branched out into health care, and I built ratings of doctors, hospitals, and other healthcare products and services. This experience helped me learn about metrics of health care quality, which qualified me for a position at 3M Health Information Systems (HIS), where I trained healthcare advisers in interpreting data and transforming it into actionable information for executives in healthcare organizations, and translated complex quantitative ideas between health economists and software engineers who were building large-scale data applications to support improvements in healthcare quality. My statistical skills were invaluable to me as a researcher in math education, allowing me to conduct quantitative research by collecting my own data, and to train graduate students in being effective consumers of research.

Each step in this path helped prepare me for the next step, and yet the journey has not been linear. You really don't need a plan about your entire career; at each step, you just need to find the next interesting thing to do.

Use data analytics and statistics to learn what data has to say. There is a common misperception about statistics and data science. Crunching numbers is an important part of the work, but designing effective research plans to collect meaningful data and interpreting that data to transform it into a useful form are all essential to applied statistics. A dataset is just a set of numbers which doesn't really mean much on its own. A statistic gets us a few steps closer to something meaningful by summarizing data, but only once the statistics are interpreted in the context of the original scientific (or other) problem do we have *information*. When that information is used to make decisions or to understand something, we have built *knowledge*.

An applied statistician is like a detective, following a trail of quantitative clues to solve an important problem. A statistician looks at data as many different ways as possible until she is confident she knows the story the data has to tell. Then, she chooses the most straightforward way to present the findings to others. Being an effective statistician or data scientist is so much more than building code and mucking around in data—statistics is the science of extracting meaning from data, and for me, that's the really fun part. Before diving into any dataset, learn where the data comes from, what it represents, and the research question

Abbe Herzig is the director of education at the American Mathematical Society. Her email address is ahh@ams.org.

[15] *As far as I can tell, the difference between statistics and data science is that the latter uses sophisticated software and database tools to apply statistical and other quantitative reasoning to today's large datasets ("big data").*

DOI: https://dx.doi.org/10.1090/noti2063

or goal of the project. Without that, you'll be limited in the insights you can find in the data. It also allows you to give feedback to collaborators about the strengths and weaknesses of the data and the types of questions it can or cannot answer. Curiosity to learn about the data and its interpretation and presentation makes a statistician valuable to a team.

Strong communication skills open lots of doors. At *Consumer Reports* and 3M HIS, I interpreted complex quantitative information for a variety of nontechnical audiences. I worked closely with engineers, scientists, writers, and editors to translate data into information that was engaging and useful for the end-user, including magazine readers, insurance company executives, scientists and engineers, administrators, and journalists. I worked hard at statistics, and even harder to learn to write well and to make effective presentations, and I developed expertise in translating data into information. This expertise allowed me to do some pretty interesting things alongside my mainstream career path, such as working as a consultant for the United Nations Statistics Office on the presentation of data about the health, economic, and educational status of women around the world, and for the Urban Institute about crime and economic mobility.

The way you might discuss quantitative information with a statistical colleague is different from how you would write a report for nontechnical consumers or for an executive of a health insurance company. There are standard practices in different disciplines about how to describe statistical work in published articles and scholarly presentations. Successful communication depends on the type of communication, and also on knowing who you are communicating with. For example, when I worked on the discrimination lawsuit I described earlier, we had to present our findings to a judge who would decide the outcome of the case. There is a complicated history of using statistics in court, and many judges and lawyers are hesitant to rely on statistical evidence if they do not clearly understand it. We were confident we had found evidence of discriminatory practices at this employer, and were committed to helping the judge understand how we drew that conclusion. Our particular challenge was how to explain statistical significance concisely to a nonexpert. Significance is all about probability, and probability is challenging for some people to understand. Our attorneys suspected that the judge was a card player, and that gave us our opening to present our results by comparing them to the probabilities of various poker hands. It worked, and the Legal Defense Fund won the case.

The ability to make the information contained within data accessible to your audience can differentiate you from other statisticians and lead to many interesting professional projects.

Concluding thoughts. Communicating about quantitative information has become the cornerstone of my statistical career, and I rarely analyze data anymore, focusing instead on training others in statistical thinking. There are many other paths you can follow in statistics. For example, some statisticians find satisfaction in a world of data and software in any number of applications. Others work in theoretical statistics, developing mathematical tools and other theoretical insights. Some statisticians move among these job sectors at different times in their careers, or work in more than one at a time. Regardless of the path you are on, with training in statistics and strong communication skills,[16] you have the flexibility to move in lots of directions.

Abbe Herzig

Credits

Author photo is courtesy of the author.

[16]*Even if you are working in a language other than your native language, you can still be an effective communicator by taking advantage of your strongest communication skills (for example, writing, speaking, or visuals). Collaborating with colleagues can also help you further refine your message to reach your audience.*