Some advice on teaching statistics to social science students

Robert Andersen
University of Toronto
Outline of Talk

1. Expectations
   - Students’ aptitudes and attitudes
   - How much can we expect?

2. My goals for my statistics courses
   - Produce competent practitioners, critical consumers, methodologists?
   - Undergraduate/graduate

3. My general approach to teaching statistics
   - What topics should be covered to accomplish these goals?
   - Demystify statistics
   - Emphasis on telling a story from the data

4. Practical aspects
   - How to structure courses
1. Social Science Students Undergraduates (1)

• Most take statistics only because it is required
• Tend to have very little math background
• Many are afraid of the course, and of you
  – Some students are even hostile
• Most want to learn the material only because they want a good grade
  – they tend to be uninterested
• It is your job to make the material interesting and relevant
• Must prepare them for courses that come afterwards and try to excite them so they want to continue
1. Social Science Students
Graduate students (12)

- More students take statistics out of curiosity, but still most are there because it is required or they know they will be disadvantaged without it
  - Very few desire to become methodologists
- Much more competitive with each other than undergraduates
  - Don’t want their peers or you to see if they’re having trouble with the material
  - Need to cultivate a culture that limits the fear and competitiveness
2. Goals for the Course

Undergraduate courses:
- Encourage students to be curious
  - Not simply a ‘hoop’ students must go through
- Relevance of statistics using real world examples
- Provide a solid understanding so that they have the foundation for more advanced courses
- Emphasize concepts, not formulas (though the latter are also important)

Graduate courses:
- Again, emphasis should be on understanding concepts
- Clearly demonstrate how methods build on each other
- First course should ensure that students will be able to critically evaluate research published in top-rated journals; second course prepares them to do competent research
3. My general approach to teaching statistics (1)

- Emphasis on **concepts** and the process of **statistical analyses**
- Provide a guide for decision-making
- Remove the ‘mystery’
  - Clearly explain the **logic** behind statistical methods
  - It is not about being ‘elitist’
  - If we are truly concerned about learning about the world, the more people who understand statistics, the better
- "**There are three kinds of lies: lies, damn lies and statistics**” Benjamin Disraeli (1804–1881)
  - Statistical analysis involves as much art as science
  - I let the students in on that ‘secret’ in order to help them avoid telling ‘lies’
3. My general approach to teaching statistics (2)

- **What is a statistical model?**
  - Emphasis on parsimony

- **Tell the right story** with the data
  - Again, statistical analysis involves both science and art
  - Description versus explanation?
  - Importance of diagnostics (e.g., clearly demonstrate how nonlinearity and outliers can make coefficients meaningless)
  - Statistical significance doesn’t matter at all if you don’t get the story right

- Emphasis on how to effectively present statistical results
3. My general approach to teaching statistics (3)

- *I discuss the limitations to statistical analysis every chance I get*
- **Problems with causal inference**
  - Observational studies versus experimental studies
  - Omitted variable bias, confounding variables, mediation / moderation
  - Measurement issues
- **Inference and the sampling distribution**
  - Students must understand that *we never know* if our sample accurately represents that population
  - If you don’t make it very clear, they don’t get the fact that we are making ‘educated guesses—they will think *p*-values represent facts that appear to come from nowhere
Practical aspects (1)

- Sequence of courses is important

**First statistics course**
- Distributions, $t$-tests, and introduce linear regression

**Third statistics course:**
- GLMs, diagnostics and how violations to model assumptions can be accommodated (esp., nonlinearity and outliers).
- Worry about statistical inference only after making sure they can describe relationships and patterns well

**Fourth statistics course:**
- Extensions to GLMs, that students are likely to encounter in the literature (robust regression, mixed models, generalized additive models, structural equation models).
Practical aspects (2)

• Combination of lectures and labs/tutorials
  – I use a data projector and the white board
    – Lecture slides are brief
    – Available to students but only after the class.
• Encourage questions at any point in the class
  – I only teach material that I confidently know
• Several evaluations, usually a couple tests and weekly assignments
  
  1. Weekly assignments:
  – Students can work in groups but everyone must do their own write up
  – Make sure datasets work!! For less advanced courses, avoid allowing students to use their own data
Practical aspects (3)

2. Tests and exams:
- Large undergraduate courses usually involve short answer and multiple choice (which often require some calculations)
  - Review class before each test
  - I read the entire exam to the students (but don’t give the answers)
- Tests for graduate courses require longer answers and more extensive calculations. Usually fewer (if any) multiple choice questions.
  - I give the students a copy of the exam the week before but they cannot bring any written materials to the exam

• Extensive and regular office hours
  - Most students, especially graduate students, benefit from one-on-one meetings