Teaching Statistics to Social Science Students

Alan Agresti

Distinguished Professor Emeritus

Department of Statistics

University of Florida, USA

QM Workshop, Oxford, June 29, 2012

Outline

- My background and perspective
- Guidelines for what the introductory course should accomplish
- Course should focus on concepts rather than watered-down mathematical statistics
- Of traditional topics, what could be eliminated or receive *less* attention?
- What should receive *more* attention?
- What should be different for the introductory statistics course for graduate students?

My Background

- Trained as a statistician, not a social scientist
- Teaching in Statistics Department both “service courses” to social scientists and other non-statisticians, and also courses for Statistics majors at BS, MS, PhD levels
- UF Stat Department, in a large state university, has
  - General introductory course (including social science students, business students) on main ideas of statistics
  - Follow-up second courses specialized to particular areas (e.g., social sciences, business require second course with main focus on multiple regression, ANOVA)
  - Graduate-level sequence of two courses for social science students
  - Advanced courses for students quite comfortable with multiple regression (e.g., multivariate statistics, categorical data analysis, longitudinal data analysis).

Qualifiers

- Social science majors are required to take separate *research methods* course from their home department. I won't discuss that course.
- Most of my teaching in past 10 years was for the graduate-level courses ($\sim$ 60 students per term, with little or no TA help); the undergraduate introductory course ($>$2000 students a term) is now handled by MS-level instructors, assisted by many graduate-student TAs.
- My comments apply to general introductory statistics courses at undergraduate level, not just those for social scientists. This partly reflects less specialization at U.S. undergraduate curriculum than in Britain.
- My opinions partly reflect how I feel about the way Statistics is presented in introductory textbooks.

The General Introductory Course: GAISE Reports

Guidelines for Assessment and Instruction in Statistics Education (GAISE) project supported by American Statistical Association created recommendations for introductory statistics courses.

See `www.amstat.org/education/gaise`

Recommendations include:

1. Emphasize statistical thinking and conceptual understanding, rather than mere learning of recipes for different methods. Statistics is a *process* to answer questions (and unlike math, perhaps no unique answer!), not a toolkit of formulas.
2. Foster active learning (e.g., activities, projects).
3. Use technology (applets, simple software) to aid conceptual understanding and reduce computational drudgery.

Emphasize statistical thinking and concepts

- Conclusions from a well designed study beat anecdotes.
- Variability, and how it is quantifiable with an appropriate study:
  - Random assignment in a controlled experiment allows cause and effect conclusions.
  - Random sampling in a survey allows us to make inferences about the population of interest.
- Limitations of studies with observational data, common sources of bias in surveys
- How associations are affected by “lurking variables” (e.g., in U.S. murder trials, proportion of defendants who get the death penalty is higher for whites than blacks, but much higher for blacks when adjust for race of victim.)
- Association does not imply causation.

Emphasize statistical thinking and concepts (2)

- Concept of a sampling distribution, and how it relates to making inferences from samples
- Significance testing and its limitations
  - *Statistical* significance does not imply *practical* significance.
  - A lack of statistical significance does not mean H_0 is true.
- Confidence intervals, and how we learn more from them than from a significance test
- Experience how to critique reports in newspapers and on Internet and journal articles that have statistical information
- Understand the processes statisticians/methodologists use when formulating and conducting research

Concepts rather than recipes

Danger: Student confusion from number of topics; e.g., for significance tests and confidence intervals for means and for proportions, we should not try to cover all combinations of:

- one sample, two sample, many samples
- univariate, multivariate response variable
- independent samples, dependent samples
- parametric (normal, binomial), nonparametric
- one-sided, two-sided
- large-sample, small-sample

If students don't understand concepts, not much gain by learning recipes for how to analyze data in various situations.

Concepts rather than recipes (2)

As put more emphasis on concepts and interpretations, can put less on formulas (except for helping to explain the concepts) that are a roadblock for students with poor algebra skills; this helps also with the “mixed ability” issue typical in such courses.

Tell students at beginning of course that if they think they’re poor at math, algebra, they can still do well in the course.

In UK, many courses are brief (e.g., 10 lectures plus 10 lab sessions), so it seems especially crucial to focus on the “big ideas” rather than math theory and formulas.

On exams, can use multiple-choice questions and written interpretations (including using software output) to focus on concepts and appropriate interpretations rather than the technical details of how to plug numbers into formulas to get certain answers.

Active learning

- Perhaps get data for examples from class survey, ideally with students having input in formulating questions of interest and developing measuring instrument.
- During second half of course, have teams of 2-3 students conduct projects, perhaps presenting results to the class on a poster accompanied by 15-minute talk.
- Perhaps analyze different aspects of an interesting data set at various points in the course. (They should look at data right from the start, rather than spending much of the course on "prerequisites" such as probability before getting to data analysis.)
- Introduce classroom activities in context of real problems.
- Good resource for classroom activities: *Activity-Based Statistics* by Scheaffer et al., Springer-Verlag (1996)

Examples of activities

- Create contingency tables using variables of interest at General Social Survey (`sda.berkeley.edu/GSS`), or find variable strongly associated with response variable identified by instructor
- Using applet to generate sampling distribution of a proportion for various n
- Why 0.05 is a common significance level (placebo better than treatment for observation 1, 2, 3, 4, 5, ...)
- Do literature search and critique article on topic of interest: Study design? Observational study or experiment? Response and explanatory variables? Statistics used? Conclusions? Limitations of study (e.g., confounding variables)? What could have been done better?

An activity I used on first day of course:

Ex. How does randomness look?

(Apparent trends, such as “hot hand” in sports, stock market up/down, may reflect mere random variability)

For n flips of coin (outcomes “head” and “tail” might represent “favor” and “oppose” in survey or “Labour” and “Conservative” in even election)

$$E(\text{longest run of heads}) \approx \text{linear in } \log(n)$$

4 for n = 25 flips, 5 for n = 50, 6 for n = 100, 7 for n = 200

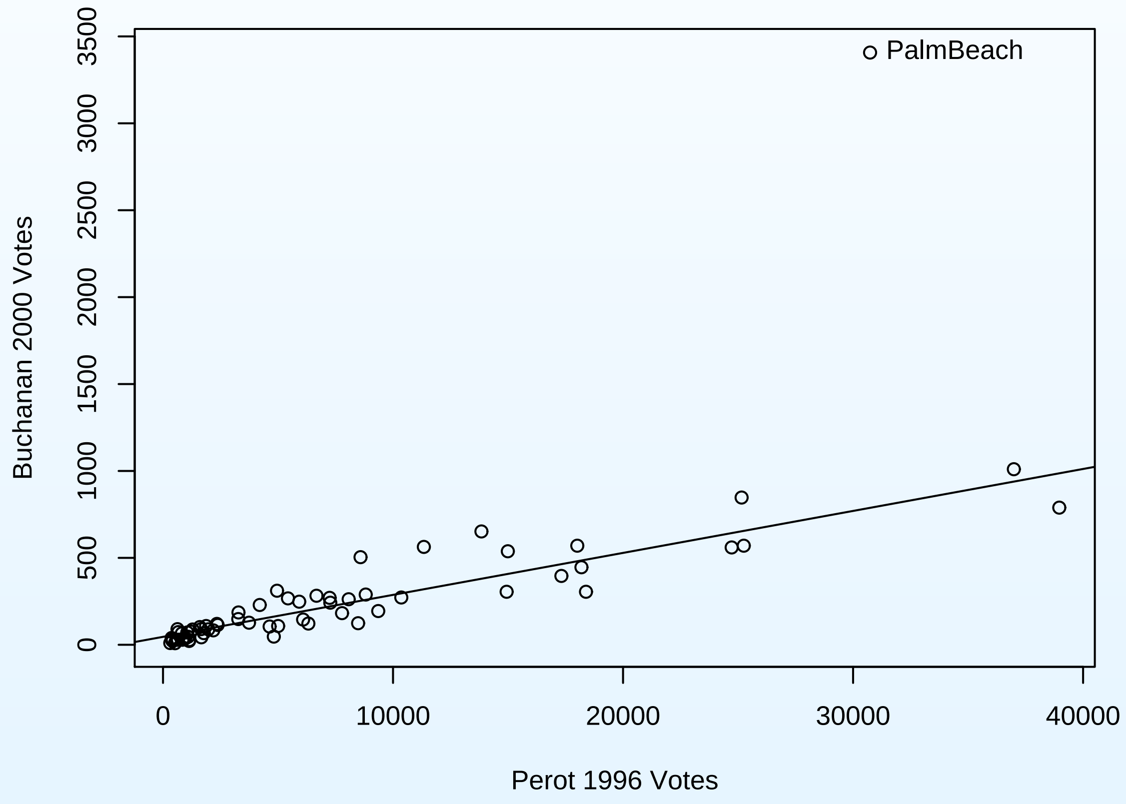
Can use to explain that randomness has unpredictable aspects but also predictable aspects that a sampling distribution describes (degree of variability in sample proportions, law of large numbers)

Technology for aiding conceptual understanding

- Use software for computations (something simple that does not take much time to teach, such as SPSS or Minitab).
- Using software (and interpreting output in examples) helps the course to focus on concepts rather than computational details of formulas.
- Use only needed formulas, and in form that enhance understanding (e.g., ignore “short-cut” formulas for variance, correlation, regression coefficients).
- Explore “what happens if ...” questions, such as showing effects of outlier on results
- Get students in habit of exploring data with graphics and basic descriptive summaries before using more complex methods, and to help check assumptions (e.g., about regression model), search for unusual observations.

Ex. Florida vote by county (Bush/Gore election):

Buchanan in 2000, Perot in 1996 (Reform party)



Technology for aiding conceptual understanding (2)

Use applets (preceded by physical explorations) to help explain difficult concepts such as sampling distributions (and effect of n on its spread), inferential error rates.

See `www.stat.tamu.edu/~west/ph/` for some good ones by Webster West that I've used in textbooks and my teaching.

- Sampling distributions:
  `www.stat.tamu.edu/~west/ph/sampledist.html`
- Significance tests:
  `www.stat.tamu.edu/~west/ph/propht.html`
- Confidence intervals:
  `www.stat.tamu.edu/~west/ph/propci.html`
- Regression by eye (and seeing effect of moving a point)
  `www.stat.tamu.edu/~west/ph/regeye.html`

Technology for aiding conceptual understanding (3)

There are now many on-line resources where statistics instructors share ideas and research about statistics education.

- Free on-line *Journal of Statistics Education*, `www.amstat.org/publications/jse`

  Website resources include a data archive.

- Free on-line *Technology Innovations in Statistics Education*, `escholarship.org/uc/uclastat_cts_tise`

  e.g., current issue: “Using applets and video instruction to foster students’ understanding of sampling variability”

- The International Association for Statistical Education, `www.stat.auckland.ac.nz/~iase` and its *Statistics Education Research Journal*.

Some things to de-emphasize (in my opinion)

- Inference with population std. dev. σ assumed known
  (can first use z instead of t in inference by doing proportion before mean)

- Less emphasis on significance tests and more on confidence intervals
  (natural to see concept of “margin of error” after learning about standard error of a sampling distribution)

- Less emphasis on one-sided tests
  (agreement with practice, confidence intervals)

- Less classical probability (e.g., counting rules, Bayes rule and complex probability calculations, continuity correction in approximating binomial with normal) and probability distributions (cover normal, skip Poisson, gamma, possibly binomial)

Some things to de-emphasize (2)

- De-emphasize strict frequentist interpretations (vs. Bayesian)
- Drop nonparametric and small-sample methods (focus on robust parametric methods, such as two-sided inferences)
- Dependent samples (e.g., longitudinal studies) can wait until a later course
- Some assumptions/conditions
- Some important ideas (e.g., probabilistic concepts) can be introduced at a lower technical level using a simplified approach

Simplify in presenting difficult concepts:

Ex. *Conditional probability* is important but difficult, can be explained in intuitive terms with contingency tables, tree diagrams.

In a criminal justice setting, suppose legal trials result in

P(conviction | guilt) = 0.9

P(acquit | innocent) = 0.9

Then, if P(guilt) = 0.1, what is P(guilt | conviction)?

Ex. Contingency table for conditional probability

Reality	Trial result		
	Convict	Acquit	Total
Guilty	9	1	10
Innocent	9	81	90
Total	18	82	100

Again, conditional on conviction, probability of guilt is about 9/(9 +9) = 0.50.

Illuminating to show students how this changes dramatically according to value of P(guilt).

Ex. Tree diagram for conditional probability

Typical results for 100 trials

```
             Guilty        Trial result
                           ----------------
             Yes (10)     |  Convict (9)
          ----------------
         |                |  Acquit (1)
         |                 ----------------
100 trials ---|
         |                 ----------------
         |     No (90)    |  Convict (9)
          ----------------
                          |  Acquit (81)
                           ----------------
```

So, conditional on conviction, prob. of guilt about 9/(9 +9) = 0.50.

What deserves more attention?: Statistics uses models

All inferential methods have *assumptions*, which form a *model* (simple approximation for reality that is never exactly true)

ex. Comparing means:

Model for pooled two-sample t test assumes

observations from group 1 at random from $N(\mu_1, \sigma)$ dist.
observations from group 2 at random from $N(\mu_2, \sigma)$ dist.

- All models are wrong (our sample was not truly random, some missing data or measurement error, distributions not exactly normal, standard deviations not identical)
- But some models are useful, summarizing lots of data by a few statistics that estimate parameters of that model
- Some assumptions, conditions much more important than others

Statistics uses models (2)

ex. Model comparing means, regression:

For two-sided inference, normal population assumption not as important as other assumptions
(unless so highly skewed that mean is inappropriate summary).

Shape of population distribution has little relevance, especially as n increases.

Data gathering assumptions important for any n.

At some point of course, we should show faulty conclusions in a published study that indicated an effect that can merely be explained by *regression toward the mean*.

Statistical model: Take recipes with a grain of salt

- Students like firm rules, but don't get too hung up on conditions, recipes, the less important assumptions.
- Students need to realize that every aspect of statistical inference involves approximation.
- Maturity in understanding statistics means getting a sense of the different sizes of approximations, which are more important.
- Statistics is the *art* and science of learning from data (and exploratory *data detective* work can lead to future useful *confirmatory* research).

What deserves more attention (2)?: Categorical data

- They'll hear about proportions the rest of their lives for results of surveys, medical research.

  (Most students will never see a std. dev., t test after this course but will see lots of proportions, margins of error and need to know when study design suggests skepticism.)

- For bivariate description, show contingency tables as well as scatterplots.

- Easier to construct sampling distributions for proportions.

- Inference for contingency tables should cover not only chi-squared; follow up by studying residuals, effect measures such as $p_1 - p_2$ and p_1/p_2.

Example: standardized residuals for contingency table

	Party Identification			
Race	Democrat	Independent	Republican	Total
White	640	783	775	2198
	(−14.2)	(2.7)	(11.9)	
Black	250	106	17	373
	(14.2)	(−2.7)	(−11.9)	

Follow up chi-squared: $X^2 = 234.7$ ($df = 2$, P-value < 0.0001)

e.g., $(17 - 114.9)/SE = -11.9$. Difference between no. of black Republicans and no. expected under H_0: independence is -11.9 SE's.

Proportion of whites who are Republican is (775/2198)$\big/$(17/373) = 7.7 times proportion of blacks who are Republicans.

Deserves attention (3): New applications/methods

Students should learn (especially in graduate-level courses) that statistics is an evolving field, not an old toolkit of recipes or a subset of mathematics

Continually new applications and methods because of modern computing power (e.g., bootstrap) and new types of data (survival analysis in medical research, data mining with huge databases such as in financial industry, bioinformatics)

What's different in introductory grad-level course?

- Most students had undergrad courses that varied greatly in quality and scope; most students remember little.
- More rapid pace
- Strong emphasis on regression (in generalized linear model terms, including logistic regression for categorical response)
- Examples and motivation more social-science research oriented (e.g., focus on journal articles rather than daily newspaper and Internet for examples)
- We're training them to be researchers, whereas we're training undergraduates to be better consumers of quantitative information.
- But it's important for student to recognize their limitations and the increasing breadth of the field, that they won't be a statistician after taking a course or two, and need to adapt an active learning strategy for their entire careers.

Textbooks focusing on a conceptual approach

Introductory statistics textbooks that changed focus from watered-down mathematical statistics to concepts first written by David Moore (Purdue), Freedman/Pisani/Purves (Berkeley)

My attempts to "do as I preach" –

Statistics: The Art and Science of Learning from Data
by A. Agresti and C. Franklin, 3rd ed. 2012

Designed for undergraduate-level general course

Statistical Methods for the Social Sciences
by A. Agresti and B. Finlay, 4th ed. 2008

1st ed. 1979 intended as student-friendly alternative to Blalock's *Social Statistics*. Power-point notes from teaching a course from this book to psychology majors at Harvard two years ago at `www.stat.ufl.edu/~aa/social/powerpoints`

Outline for intro. stats course, from *Art and Science ...*

Exploring data with graphs and numerical summaries

Association: contingency, correlation, and regression

Gathering data

Probability distributions (normal)

Sampling distributions

Statistical inference: confidence intervals

Statistical inference: significance tests

Comparing two groups

Analyzing association between categorical variables

Analyzing association between quantitative variables (regression analysis)

Thanks

THANKS for your time today and for listening to the perspective of a statistician, one who is now somewhat removed from the real difficulties you must all face!

I look forward to the discussion.