Statistics Past, Present & Future

Visit of Sir David Cox

Biostatistics Department

Johns Hopkins Bloomberg School of Public Health

June 25, 2008

Purpose

Free-wheeling discussion

Some references

- Bartholomew DJ. What is statistics? *JRSSA* 1995; 158:1-20.

- Cox DR. Applied statistics: A review. *Ann Appl Stat* 2007; 1:1-16.

- Demets DL, … Louis TA, et al. Training the next generation of biostatisticians: A call to action in the U.S. *Stat Med* 2006; 25:3415-29.

- Efron B. The future of statistics. White paper, 2007.

- Lindsay BG, Kettenring J, Siegmund DO. A report on the future of statistics (with discussion). *Statist Sci* 2004; 3-387-413.

- Zelen M. Biostatisticians, biostatistical science, and the future. *Stat Med* 2006; 25:3409-14.

Why now?

- It's been on our minds
  - Department
  - Field
- A changing ecology?
  - Information age
  - "Big science"
  - Society
- Generational shift?
- Opportunity of this visit

A few views

- *"… Statistical Science [is] the particular aspect of human progress which gives the 20th century its special character…. It is to the statistician that the present age turns for what is most essential in all its more important activities."*

   Sir R. A. Fisher, 1952 address to IBS

A few views

- *"… why do so many people, who suppose that they do the very things that we do, prefer **not** to call themselves **statisticians** but operational researchers, software engineers, QA experts, forecasters or astrologers?"*

RSS News, Agent Provacateur, 1993

A few views

- "*This is the information age, statistics is the prime information science, and there is every reason to believe in a greatly increased statistical presence in the academy of the future. Or maybe not. Ideas are the coin of the realm in the intellectual world. Our continued growth and influence depends on the same thing that powered the last century, the continued production of useful new ideas and techniques.*"

  Professor Bradley Efron, 2007

"Statistics"—two connotations

- Discipline
  - NSF – Two parts: "Core"—subset of statistical activity focused inward, and the rest

- Profession
  - *"How individuals work most fruitfully in what" (DRC)*
  - Education
  - *"Statisticians have… something to contribute at the highest level[s]… which is often not recognized either among ourselves or others" (DJB)*

Future of the Field
Discussion Questions

- What will the Statistics field look like 20 years from now?

  – What does history tell us? In what ways are the prospects so novel that we're in virgin territory?

  – What are the most exciting areas of scholarship?

  – How (if at all) is the balance changing between the field's disciplinary core and specialization / interdisciplinary collaboration?

Future of statistics: Some opinions

- Highlighted disciplinary futures:
  - Data: technology revolution; massive size (BE, NSF)
  - "Compromise" methods (BE / NSF)
  - "Errors of the third kind": model selection (BE)
  - Modeling complex systems (NSF)
  - Dealing with uncertainty (NSF)
  - Middle ground between proof and computation (NSF)
  - Analysis techniques arising outside statistics (NSF)
  - Trends: more general vs. more specific models (DRC)
  - Design (DRC--?)
  - Causality (DRC--?)

Future of statistics: Some opinions

- NSF comment: *"… the most important advances will be unpredictable. For this reason we need to maintain an underlying philosophy that is flexible enough to adapt to change. At the same time is important that this future research not degenerate into a disparate collection of techniques."*

- BE comment: "*Statistics is in a period of rapid expansion and … change. During such times it pays to concentrate on basics and not tie oneself too closely to any one technology or analysis fad*."

- DRC comment: *"[though] certainly open to dispute, … the broad approaches desirable to analysis and interpretation have not been radically changed by the capacity to handle very large amounts of data, however much approaches to implementation have been and are being revolutionized."*

- DJB comment: *Our view of statistics is too small. Let us seek to broaden it.*

Future of the Field
Discussion Questions

- ## What is statistics?
  - *(Pearson, per DJB; Kendall, 1950):* *"The study of collective characters in populations."*

  - *(Grenander & Miller, 1994; NSF): The science of learning from data.*

  - *(Stevens, 1968): … "a straightforward discipline designed to amplify the power of common sense in the discernment of order amid uncertainty."*

  - *(DRC/NSF/DJB): "The discipline concerned with the study of variability, with the study of uncertainty, and with the study of decision making in the face of uncertainty."*

Future of the Field
Discussion Questions

- Some education issues
  - Decline of young people entering the field (NSF)
  - Offering sufficient depth over a wide range of tools (NSF)
  - Postdoctoral training (NSF; Demets)
  - Non-traditional areas: consulting, ethics, communication, leadership, management (Demets)
  - Balance between core & Science training (Demets)

Future of the Field
Discussion Questions

- How should we be training our students for the good of their careers and the field as a whole?

  – What innovations in content or delivery of our graduate programs are needed so that students are ready for the 10-year future?

  – In what balance should statistics students be trained as generalists or specialists?

  – What is the core of what our students should learn in their courses?

Future of the Field
Discussion Questions

- How can / should statisticians impact on science and society over the coming decades?

  – Increasing demand for statistical collaboration in science (NSF)

  – What organizational structures and rules of engagement will allow our field to flourish in what many see as a changed and changing organizational environment? (TAL)

  – What needs to be done to increase the degree to which statistics contributes at the highest levels (DJB)?