Ethics and statistics

Andrew Gelman
Department of Statistics and Department of Political Science
Columbia University

9 Apr 2014
What is an “ethics problem”? 

You are considering an action that . . .

- benefits you or some cause you support
- hurts or reduces benefits to others
- violates some rule.
Example

- A company gives you $10,000 to assist in research with a new drug, with a promise of $100,000 more if it is successful.
- But the data are inconclusive: 20/100 deaths with the treatment, 21/102 deaths with the control
- Should you . . .
  - look deeper for evidence that the new drug is better?
  - do an analysis you suspect is wrong?
  - do an analysis you know is wrong?
  - fake the data?
Marc Hauser Resigns From Harvard

By Tom Bartlett

Marc D. Hauser, the Harvard psychology professor found responsible for eight counts of misconduct by the university, has resolved pending speculation about whether the embattled professor would return to Harvard this fall.

In a letter dated July 7, Mr. Hauser notified Michael D. Smith, Harvard's dean of the faculty, of his resignation.

He wouldn't share his videotaped data...
Tough calls

- Fake “ethical dilemmas”
  - The polluting factory
- Real ethical dilemmas
  - Allocation of scarce resources
  - When to gather data and when to approve new treatments
- Ethics includes process, intermediate outcomes, and long-term outcomes
- *Any* ethics violation can be made ambiguous
- This does not negate the importance of ethics
- Statisticians should be able to appreciate the necessity of decision-making under uncertainty and ambiguity

Andrew Gelman
Ethics and statistics
Background: health effects of electromagnetic fields

Problems with the Blackman et al. study

- Use of statistical significance to categorize effects
- Noise added in the form of a comparison to an extra control group

But they didn’t know better!

But ... an ethics violation ... they refused to share their data
Some ethical problems involving uncertainty

- Cost-benefit analysis for environmental hazards
  - Problem with “zero tolerance” for risk
  - Problem with no regulation
- Medical statistics:
  - The impossibility of equipoise
  - Informed consent and randomization
  - Long-term benefits of a new treatment
The Commissar for Traffic presents the latest Five-Year Plan.
Thinking about ethics

- Laziness
- Lying, cheating, and stealing
- Mistakes such as selection bias
- Systematic remedies (the replication and criticism movement)
- Tradeoffs regarding uncertainty
- Is it possible to be an ethicist without being mean to people?
Are mainstream statistical methods themselves unethical?

- P-values and statistical alchemy
- “The AAA tranche of subprime science”
- Bayesian inference and subjective science
- The statistical significance filter
- Incentives to exaggerate
The estimated change in life expectancy (and height of the brace) just north of the Huai River is -5.04 years and is statistically significant (95% CI: -8.81, -1.27).

**Fig. 3.** The plotted line reports the fitted values from a regression of life expectancy on a cubic in latitude using the sample of DSP locations, weighted by the population at each location.
Some writings on ethics by statisticians

Some other references on ethics


▶ much more . . .