Statistical Ethics

G. Oehlert & S. Weisberg

School of Statistics
University of Minnesota

May 4, 2009

Statistics and the Law

Is there objective, legal truth?

Ref: P. Meier (1986), "Damned Liars and Expert Witnesses", *J. Amer. Statist. Assoc.*, 81, 269-276.

Sales tax has been erroneously applied to items not subject to tax. Plaintiff seeks to estimate the amount of the tax applied. Available are sales slips for 826 business days; all sales slips must be examined by hand to determine amount, a very expensive and potentially error-prone proposition.

As statistician for the *plaintiff*, what do you do?

If you work for the defense

Suppose sampling had been done imperfectly by the other side, for example, using some systematic method. What is your role?

. . . if you believe sampling was not representative?

. . . if you believe that the bias caused by poor sampling is minor?

. . . if the sampling for the other side was done by "Dr. Deming or Professor Cochran"?

From Meier, p. 270

". . . Cochran was fond of telling of the occasion when he was called on to carry out a sampling study of . . . retail stores, and he instructed that the sample consist of every tenth establishment . . . in the Yellow Pages. The judge, he said, welcomed his expert testimony as a learning experience and remarked, after Cochran has been swarn, 'I am glad to hear and learn from Professor Cochran about this scientific sampling business, because I know virtually nothing about it. In fact, about the only thing I *do* know is that you should not just start at the beginning and take every 10th after that.'"

Moral: Be overly cautious.

It's the law

SD rule "... as a general rule ... if the difference between the expected value and the observed number is greater than 2 or 3 standard deviations, then the hypothesis that the jury drawing [the issue in this case] was random would be suspect to a social scientist" *Casteneda v. Partida*

Preponderance of the evidence In a civil non-criminal lawsuit, decisions are based on the more convincing evidence and its probable truth or accuracy, and not on the amount of evidence. This contrasts with "beyond a reasonable doubt" in criminal trials. Thus, for example, a Bayes factor of 1.01 could be enough. *Source: dictionary.law.com*

Corrupting influences, also from Meier

1. "If you find the expert you choose is independent and not firmly committed to your theory of the case, be cautious about putting him on the stand. Your cannot be sure of his answers on cross-examination. when I put an expert on the stand, he is going to know which side we are on." (John C. Sheperd, cited in Meier, p. 273)
2. "[T]he statistician is tempted to give the definitive rather than a qualified answer to key questions ... [and] to ignore or minimize qualifications that he might emphasize in a more academic setting..." (Meier, p. 273)
3. Bribery
4. Flattery
5. Co-option to agree with the cause (Stockholm syndrome).
6. Gladiatorial role.
7. Personal views.

Cheating

Source: Peggy Fischer, NSF Office of Inspector General, `http://www.nsf.gov/oig/BMCCResearchIntegrityandPlagiarism.ppt`, Jan 2008:

- 75% of high school students admitted to cheating
- 53% of undergraduates admitted to cheating
- 30% of researchers admitted to "questionable practices"
- NSF has observed a doubling of allegations and increasing numbers of serious cases resulting in research misconduct proceedings.

Plagiarism

Many forms:

- Word for word copying without quotation or citation.
- Stitching together random bits from various sources without careful identification.
- Paraphrase without acknowledgment of the original.
- Casual repetition of great term or phrase.

Martin, Ohrman, Wheatley.

Plagiarism everywhere?

Many examples in science, academia, and journalism, e.g., Jayson Blair at the *New York Times*, who among other things copied a report from another paper and submitted it has his own.`http://en.wikipedia.org/wiki/Jayson_Blair`.

Type "Reports of plagiarism" into Google: First result is a U of MN form.

Very easy to do in a click-and-drag internet age.

However, only about 20 cases a year get reported to NSF and NIH.

Chronicles of Higher Ed., Feb. 22, 2008

"A Columbia University professor who was found to have committed numerous acts of plagiarism struck back at her accusers on Thursday, saying it was they who stole her work and accusing administrators of blackmail and intimidation."

The professor was found guilty of appropriating ideas and even text from work of students in her department. The professor published before the students' work was published. The professor claims that students stole her work, although the findings of a hearing were that she stole from her students.

http://chronicle.com/daily/2008/02/1798n.htm?utm_source=at&utm_medium=en

Intellectual Property a la Minnesota

> *Any invention, discovery, improvement, copyrightable work, integrated circuit mask work, trademark, trade secret, and licensable know-how and related rights. Intellectual property includes but is not limited to, individual or multimedia works of art or music, records of confidential information generated or maintained by the University, data texts, instructional materials, tests, bibliographies, research findings, organisms, cells, viruses, DNA sequences, other biological materials, probes, crystallographic coordinates, plant lines, chemical compounds, and theses. ...*

Intellectual Property a la Minnesota

> *... Intellectual property may exist in a written or electronic form, may be raw or derived, and may be in the form of text, multimedia, computer programs, spreadsheets, formatted fields in records or forms within files, databases, graphics, digital images, video and audio recordings, live video or audio broadcasts, performances, two or three-dimensional works of art, musical compositions, executions of processes, film, film strips, slides, charts, transparencies, other visual aid/aural aids or CD-ROMS.*

University of Minnesota Intellectual Property Policy

Types of IP

Copyright A copyright protects the tangible expression of an idea, not the idea itself (e.g., a book, a research article, or a videotape).

Patent A patent protects the idea and gives the creator the right to exclude others from using the idea. To receive a patent, the creator must disclose in detail how to make his invention work and its use.

Trademark A trademark identifies and distinguishes an idea, written words, pictures, or products from those of competitors (e.g., golden arches).

Trade Secret A trade secret refers to information that is not publicly known, that produces economic benefit to the owner, and that the owner maintains as secret.

No Claim on Academic work

University does not claim ownership rights to

> *Intellectual property created by a student for the sole purpose of satisfying course requirements unless the student assigns ownership rights in the intellectual property to the University in writing or assignment of such ownership rights to the University is made a condition of participation in a course.*

No Claim on Academic work

University does not claim ownership rights to

> *Regular academic work product: any copyrightable work product that is an artistic creation or that constitutes, or is intended to disseminate the results of academic research or scholarly study. Regular academic work product includes, but is not limited to, books, class notes, theses and dissertations, course materials designed for the web, distance education, and other technology-oriented educational materials, articles, poems, musical works, dramatic works, pantomimes and choreographic works, pictorial, graphic and sculptural works, or other works of artistic imagination. Software specifically needed to support a regular academic work product or that is designed to disseminate the results of academic research and scholarly study is also considered a regular academic work product.*

Otherwise, they own you

> *The University owns all intellectual property created through the use of University resources or facilities, supported directly or indirectly by funds administered by the University, developed within the scope of employment by employees, assigned in writing to the University, or agreed in writing to be a specifically commissioned work.*
> *In the case of intellectual property created in the course of sponsored research or under contract with external parties, ownership is determined in accordance with the terms of the University's agreement with the funding agency or external party and applicable law. The same applies to intellectual property created under outside consulting or service arrangements.*

Who owns what?

1. Professor KnowItAll presents oral lectures, and Student Writesitup copies them down. Student is the owner.
2. Professor KnowItAll reads his lectures, and Student Writesitup copies them down. KnowItAll is the owner.
3. Student writes a paper in a class. Student is the owner.
4. Student participates in a class blog. Who owns the student's remarks in the blog? Student is the owner.
5. Student's review of an article is posted on a class website. The following year, this post is required reading in a course at a different university. Is this a violation of copyright? Yes, but academic fair-use is a good defense (Senate Res. Comm minutes, 9/17/07)

Other employment

Many employment contracts, including consulting, contain IP ownership clauses.

Read and check the terms of employment carefully!

If you are a programmer for a drug company and you write the next great statistics package at home in your spare time, the drug company may still own it.

Do not try to sell what legally belongs to your employer!

Disclose IP to its owner.

Conflict of interest

A conflict of interest occurs when Academic Employee compromises his/her professional judgment in carrying out University teaching, research, outreach, or public service activities because of an external relationship that directly or indirectly affects the Financial Interest of the Academic Employee, and Family Member, or any Associated Entity.

Examples

- A major interest in a private firm by a faculty member who also has the decision-making responsibility in awarding a contract to that firm.
- Sponsorship of research by commercial firms in which the faculty member has a significant interest.
- Nepotism.

More Conflicts

Scientific Participation in review/referee panels regarding the allocation of resources or the publication of papers. Usually handled by excusing the person with the potential conflict.

Academic Utilization of the name and/or the resources of the University for personal gain.

Commitment Spending too much time on non-University activities (consulting, service, etc) leading to a significant decrease in the time and effort devoted to the employer.

What to do?

Disclosure, Disclosure, Disclosure.

All potential conflicts must be disclosed to appropriate parties (e.g., supervisors).

There may be no problem, but make sure in advance.

Data management

Data take many forms:

- measurements
- images
- interviews
- recorded behaviors
- medical records
- school records
- physical artifacts
- etc

Data topics

- Reliability
  - Records should indicate what, why, who, when.
- Maintenance
  - Records should be permanent and err on the side of thoroughness.
- Retention
  - Principal investigator in charge of keeping data
  - But ...
- Access
  - Protect private records
  - Satisfy contractual requirements
  - May be subject to litigation

Human Subjects

Principles:

- Respect
- Beneficence
- Justice

which generally requires

- Informed consent
- Reasonable risk/benefit ratio
- Equitable selection of treatments

Human (and Animal) Subjects, continued

Research design must

- Balance benefits against harm
- Use efficient, correctly sized designs
- Maximize information extracted from subjects
- Consider interim analysis
- Avoid placebo-controlled testing when the condition is harmful
- Respect privacy (HIPPA: It's the law!)

Institutional Review Board

Federal funds & human subjects $\longrightarrow$ IRB

IRB protects rights and welfare of human subjects.

IRB must approve any project with human subjects.

Non-profits (e.g., Universities) generally have their own volunteer IRBs; other research organizations may use professional IRBs.

Investigator must

- Possess skills needed for the research
- Design research that meets risk/beneift criteria
- Submit research plans to IRB
- Ensure prior informed consent
- Protect vulnerable populations
- Train personnel
- Adhere to high ethical standards
- Ensure privacy
- Keep records/make reports
- Comply with all regulations

IRB will

- Weigh risks and benefits
- Deny substandard designs
- Provide oversight through completion

but...

- Researchers don't like IRB's.
- IRB's can ask for the impossible.
- IRB's can be very slow

Social Responsibility

As a scientist

- To keep an open mind. Report honestly what you see.
- To seek evidence and not accept things at face value.
- To exercise due diligence and not be sloppy.
- To make full disclosure to the extent possible.
- To stay within the limits of your expertise.
- Don't over interpret.

Social Responsibility

As a member of society

- To use resources wisely.
- To credit sources and acknowledge priority.
- To deal fairly with colleagues and subordinates.
- To act on ethical breaches by others.

Ethics for statisticians

Finally, we now look at some ethical ideas that target statistics a little more closely.

Sometimes statistics is science, so we need scientific ethics.

Sometimes statistics is a business, so we need business ethics

Sometimes statistics is societal, so we need social ethics.

ASA Guidelines

We begin with the American Statistical Association's *Ethical Guidelines for Statistical Practice*, which can be found at
`www.amstat.org/profession/index.cfm?fuseaction=ethicalstatistics`
These guidelines list eight areas to consider.

Professionalism

1. Guard against predisposition of results
2. Keep current in methodology
3. Acquire adequate statistical and subject matter expertise
4. Work with ethical collaborators
5. Assure appropriateness of procedures (even automated ones)
6. Deal with multiple comparisons
7. Respect contributions and IP of others
8. Disclose conflicts of interest

Funders, Clients, Employers

1. When possible, present a choice among valid alternatives
2. Clearly state your qualifications and experience
3. Clarify roles in a project
4. Don't work towards predetermined outcome
5. Fulfill commitments
6. Accept responsibility for your performance

Publication and Testimony

1. Maintain personal responsibility for work bearing your name
2. Report assumptions and question them
3. Make it clear who did statistical work (pubs or testimony)
4. Have clear basis for authorship order
5. Account for all data considered
6. Report sources and adequacy of data
7. Report/document data cleaning/imputation procedures
8. Address potential confounding variables
9. Identify financial sponsor
10. Report limits on inference and sources of error
11. Share data (modulo privacy/confidentiality)
12. Promptly and publicly correct errors
13. Write to the audience

Official Statistics

1. "...unaddressed ethical issues simply threaten the credibility of a statistical agency, undermining the trust that data providers, the public, and policy makers place in the agency or government statistical work generally."
2. Use sound methodology.
3. Privacy *vs* collective needs (homeland security, for example)
4. Conflict between politics and science
5. Equal and public access to data
6. Work to maintain the integrity of statistical agencies and national statistical offices. Without good data, can we expect good decisions?
7. "Biometrics".

Source: W. Seltzer (2005). Official statistics and statistical ethics: Selected Issues. ISI proceedings.

Misconduct

1. Do not condone careless, incompetent, or unethical statistical practices
2. Deplore professional misconduct
3. Differences of opinion and honest error are not misconduct
4. Follow procedures in a misconduct investigation
5. Help pick up pieces after misconduct investigation
6. Do not retaliate against whistle blowers