

Catalog Title: Ethical Practice of Statistics
Statistics, Ethics, and the Social and
Behavioral Sciences
Psychology/Statistics 484

Course Synopsis: This course is devoted to the ethical practice of statistics, defined as being in accord with the accepted rules and standards for right conduct that govern the discipline of statistics and its many areas of application. An emphasis is placed on the use of statistical and probabilistic reasoning in the social, behavioral, and biomedical sciences, with particular stress on the relation to law and the judiciary. Among others, topics include the use of Bayes Rule in screening for rare events; the importance of base rates; probability of causation; inferring causality; meta-analysis; ethics in data collection; the Federal Rules of Evidence; experimental design and data interpretation; Simpson's Paradox; statistical sleuthing and explanation; observational and controlled studies; the social construction of statistics; prediction and correlation.

Primary Text: *A Statistical Guide for the Ethically Perplexed* (Acronym: SGEP), Lawrence Hubert and Howard Wainer, Chapman & Hall, 2013. This text is available as an e-book through our library (under a CIC publication agreement with Taylor and Francis).

<http://hdl.handle.net/2142/42035>

The thirty (required) longer readings from the popular literature

(e.g., *New Yorker*, *New York Review of Books*) and the Appendix Supplements and other suggested readings are at:

http://cda.psych.uiuc.edu/sgep_course_material

this syllabus: [sgep_course_syllabus.pdf](#)

required (longer) readings: [sgep_readings_syllabus_ordered.pdf](#)

appendix supplements and lists of suggested readings:

[sgep_supplements.pdf](#)

Also, various slides and handouts used in class will be placed at this site in due course.

Prerequisites: An introductory statistics class, e.g., PSYC 235, PSYC 301, STAT 100, ECON 202, EPSY 480, SOC 280.

Audience: upper-division undergraduates (particularly those who are pre-law, pre-med, or majoring in statistics); first-year graduate students in any of the social and behavioral sciences.

Grading: Grades will be based on three one-hour exams given in class plus the final exam.

Exam 1: given in week 5; covers first four weeks

Exam 2: given in week 9; covers second four weeks

Exam 3: given in week 13; covers third four weeks

The final exam is cumulative but with some greater emphasis on the last three weeks.

One 8 1/2 by 11 handwritten “crib” sheet that students prepare will be allowed for each one-hour exam (both the front and back of the sheets can be used); these will presumably contain notes on the readings and lectures, definitions, formulas that might be needed, and

so on. The final examination will allow the three “crib” sheets from the one-hour exams plus one more covering the last three weeks. Exam questions will vary from multiple choice, true/false, fill-in a word or phrase, short answer, short computational problems (so, a hand-calculator is required), matching.

There is no “curve” for this class. To determine your final grade, add up your four exam scores (each worth a maximum of 40 points). The following chart indicates how final letter grades will be assigned:

A+ : 155–160; A : 149–154; A- : 143–148; B+ : 137–142; B : 131–136; B- : 125–130; C+ : 119–124; C : 113–118; C- : 107–112; D+ : 101–106; D : 95–100; D- : 89–94; F: 88 or below

No makeup exams will be given. If you miss an hourly exam (a maximum of one), your grade will be based on the other three scores with appropriate modification of the grade ranges.

For graduate credit: an additional paper is required covering some substantive topic needing your original and cogent statistical/probabilistic thought (about 15 pages); this paper must be graded as “satisfactory.” For representative topics that might be of interest, see the required and suggested readings in the syllabus; also, search for relevant statistical articles/applications in the behavioral sciences on the *New York Times* website: try, for example, “Benedict Carey” or “Gina Kolata”. There is a file on the class web site (projects_sgep.pdf) that lists a few suggestions and sources for possible class projects. (If you wish to do one of the topics numbered 1 to 11, let me know quickly [say, by the end of the second week of classes]; I may have to do a lottery if multiple people wish to do the same specific topic.) A short paper proposal is due in week 8; the complete paper itself is due by

the final examination date. As an alternative form to a written paper, the class project could be in the form of slides (and associated notes) for a “brown-bag” presentation of about one hour in length.

Graduate students might wish to use this opportunity to generate a publication for an outlet such as *Significance*, the joint (popular) statistics venture between the Royal Statistical Society and the American Statistical Association. See

www.significancemagazine.org

for the type of article they publish (all issues of *Significance* are available on-line through the library).

Academic integrity: Students are reminded of the University’s policy on academic integrity as it pertains to conduct in examinations and in the writing of papers:

http://admin.illinois.edu/policy/code/article1_part4_1-401.html

This statement reads in part:

The University has the responsibility for maintaining academic integrity so as to protect the quality of education and research on our campus and to protect those who depend upon our integrity. It is the responsibility of the student to refrain from infractions of academic integrity, from conduct that may lead to suspicion of such infractions, and from conduct that aids others in such infractions

Initial Contact Information: Instructor: Lawrence Hubert; Office: 424 Psychology (lhubert@illinois.edu) Class meetings: Tuesday and Thursday, 3:30–5:30, Psychology 23.

Semester Schedule

(For synopses of the various films to be shown [and several others that won't], see the file on our web site: **films_sgep.pdf**)

(For lists of keywords sorted by each week's required reading, see the file on our web site: **keywords_sgep.pdf**)

Week 1: Introduction to (and Review of) the Course

Required Reading:

SGEP (vii–xiv; 1–15) —

0. Preface

1. Preamble

2. Introduction

2.1 The (Questionable) Use of Statistical Models

Popular Articles —

Why I Wrote “The Crucible,” Arthur Miller (*New Yorker*, October 21, 1996)

The Trauma Trap, Frederick C. Crews (*New York Review of Books*, March 11, 2004)

Health Care: Who Knows ‘Best’? Jerome Groopman (*New York Review of Books*, February 11, 2010)

Suggested Reading:

1.1 Preamble: Additional Suggested Reading

Film: *The McCarthy Years* (114 minutes)

Week 2: Probability Theory: Background and Bayes Theorem

— the case of Sally Clark, wrongly convicted in England of killing her two children; this miscarriage of justice was due to an inappropriate assumption of statistical independence and the commission of the “Prosecutor’s Fallacy”

— breast cancer screening through mammograms; understanding Bayes’ theorem, test sensitivity and specificity, prior probabilities, and the positive predictive value (for example, what is the probability of having breast cancer if the mammogram is “positive”?)

Required Reading:

SGEP (19–86) —

3.1 The (Mis)assignment of Probabilities

3.2 The Probabilistic Generalizations of Logical Fallacies Are No Longer Fallacies

3.3 Using Bayes’ Rule to Assess the Consequences of Screening for Rare Events

3.3.1 Ethical issues in medical screening

3.4 Bayes’ Rule and the Confusion of Conditional Probabilities

3.5 Bayes’ Rule and the Importance of Base Rates

3.5.1 The (legal) status of the use of base rates

3.5.2 Forensic evidence generally

Popular Articles —

Trawling the Brain, Laura Sanders (*ScienceNews*, December 19, 2009)

The Cancer-Cluster Myth, Atul Gawande (*New Yorker*, October 4, 1999)

Duped, Margaret Talbot (*New Yorker*, July 2, 2007)

Suggested Reading:

- 3.1 Suggested Reading Relevant to the Whole Chapter
- 3.2 Suggested Reading on Probability Issues
- 3.3 Suggested Reading on Forensic Issues
- 3.4 Suggested Reading on Screening

Films: *Unfinished Business* (58 minutes); *The Infamous Dreyfus Affair* (50 minutes)

Week 3: Probability Theory: Application Areas

— how subjective probabilities might be related to the four levels of a “legal burden of proof”: “preponderance of the evidence”; “clear and convincing evidence”; “clear, unequivocal, and convincing evidence”; and “proof beyond a reasonable doubt”

— the distinction between “general causation” and “specific causation”; the common legal standard for arguing specific causation as an “attributable proportion of risk” of 50% or more

— issues of probability, risk, and gambling; spread betting and point shaving; parimutuel betting; the importance of context and framing in risky choice and decision-making

Required Reading:

SGEP (87–118) —

4.1 Some Probability Considerations in Discrimination and Classification

4.2 Probability and Litigation

4.2.1 Probability of causation

4.2.2 Probability scales and rulers

4.2.3 The cases of Vincent Gigante and Agent Orange

4.3 Betting, Gaming, and Risk

4.3.1 Spread betting

4.3.2 Parimutuel betting

4.3.3 Some psychological considerations in gambling

Popular Articles —

Better Decisions Through Science, John A. Swets, Robyn M. Dawes, and John Monahan (*Scientific American*, October 2000)

Do Fingerprints Lie? Michael Specter (*New Yorker*, May 27, 2002)

Under Suspicion, Atul Gawande (*New Yorker*, January 8, 2001)

Suggested Reading:

4.1 Suggested Reading on Agent Orange and Judge Weinstein

4.1.1 Appendix: The Redacted Text of Judge Weinstein's Opinion in the Fatico Case

4.1.2 Appendix: Guidelines for Determining the Probability of Causation and Methods for Radiation Dose Reconstruction Under the Employees Occupational Illness Compensation Program Act of 2000

4.1.3 Appendix: District of Columbia Court of Appeals, In Re As. H (Decided: June 10, 2004)

4.2 Suggested Reading on Issues of Risk

4.3 Suggested Reading on Issues of Betting and Gaming

Film: *The Central Park Five* (2 hours)

Week 4: Correlation

— “Voodoo Correlations in Social Neuroscience”; the “culling” or search for results in clinical trials and elsewhere, with a subsequent failure to cross-validate what is found; more generally, the problem of “double dipping”

Required Reading:

SGEP (119–140) —

5.1 Illusory Correlation

5.2 Ecological Correlation

5.3 Restriction of Range for Correlations

5.4 Odd Correlations

5.5 Measures of Nonlinear Association

5.6 Intraclass Correlation

Film: *Florence Nightingale* (65 minutes); *Snow* (22 minutes)

Week 5: Prediction

** Exam 1 given on Thursday (covers the first four weeks) **

— Henry A. Wallace and the modeling of expert judgment (“What Is In the Corn Judge’s Mind?”); the distinction between actuarial and clinical prediction, and the Dawes notion of the “robust beauty of improper linear models”

— Barefoot v. Estelle (1983): There is no merit to petitioner’s argument that psychiatrists, individually and as a group, are incompetent to predict with an acceptable degree of reliability that a

particular criminal will commit other crimes in the future, and so represent a danger to the community.

Required Reading:

SGEP (141–173) —

6.1 Regression Toward the Mean

6.2 Actuarial Versus Clinical Prediction

6.3 Incorporating Reliability Corrections in Prediction

6.4 Differential Prediction Effects in Selection

6.5 Interpreting and Making Inferences From Regression Weights

6.6 The (Un)reliability of Clinical Prediction

Suggested Reading:

6.0.1 Appendix: Continuation of the American Psychiatric Association, Amicus Curiae Brief: *Barefoot v. Estelle*

6.0.2 Appendix: Opinion and Dissent in the U.S. Supreme Court, *Barefoot v. Estelle* (Decided, July 6, 1983)

Film: *The Thin Blue Line* (102 minutes)

Week 6: The Basic Sampling Model and Associated Topics

— the decennial problem posed by the U.S. census; complete enumeration, as required by the Constitution, versus sampling, plus the political issues involved in the problem of “undercount”

— the unfortunate state of the statistical routines present in the widely used Excel program, and the inability of Microsoft to correct errors pointed out by the statistical community

Required Reading:

SGEP (175–225) —

7.0.1 Complete enumeration versus sampling in the Census

7.1 Multivariable Systems

7.1.1 Multivariable systems and unidimensional rankings

7.2 Graphical Presentation

7.3 Problems With Multiple Testing

7.4 Issues in Repeated-Measures Analyses

7.5 Matching and Blocking

7.6 Randomization and Permutation Tests

7.7 Pitfalls of Software Implementations

7.7.1 The unfortunate case of Excel

7.8 Sample Size Selection

7.8.1 Are large clinical trials in rapidly lethal diseases usually unethical?

Film: *The Plow that Broke the Plains* (27 minutes); *The River* (31 minutes); George Stoney Commentary (21 minutes)

Week 7: Psychometrics

— the settlement between the Educational Testing Services and the Golden Rule Insurance Company, and its devastating consequences for carrying out psychometrically defensible “high-stakes” testing

— the delay (because of the American Psychological Association) in the publication of the article, “Is Criminal Behavior a Central Component of Psychopathy?”; how construct validation should be

done, and the need to define a construct by more than just the Hare Psychopathy Checklist

— the darker side of psychometrics and statistics: eugenics, forced sterilization, immigration restriction, racial purity laws

— *Buck v. Bell* (1927): The Court upheld a statute instituting compulsory sterilization of the unfit “for the protection and health of the state.”

— *Loving v. Virginia* (1967): The Court declared Virginia’s anti-miscegenation statute, the Racial Integrity Act of Virginia (1924), unconstitutional, thereby ending all race-based legal restriction on marriage in the United States.

Required Reading:

SGEP (227-256) —

8.1 Traditional True Score Theory Concepts of Reliability and Validity

8.2 Test Fairness

8.3 Quotidian Psychometric Insights

8.4 Psychometrics, Eugenics, and Immigration Restriction

8.4.1 Eugenics

8.4.2 Immigration Act of 1924

8.4.3 Racial Purity Laws

Popular Articles —

Annals of Medicine: The Dictionary of Disorder, Alix Spiegel (*New Yorker*, January 3, 2005)

Personality Plus, Malcolm Gladwell (*New Yorker*, September 20, 2004)

Suggested Reading:

8.1 Suggested Reading on Psychometric Issues

8.1.1 Appendix: Excerpts From Brigham’s *A Study of American Intelligence*

8.1.2 Appendix: Racial Integrity Act of 1924 (State of Virginia); Loving v. Virginia

Film: *The Loving Story* (77 minutes); *Nazi Medicine* (54 minutes)

Week 8: Background: Data Presentation and Interpretation

— weight-of-the-evidence argumentation in presenting and interpreting data, particularly for medical and regulatory issues

— Brown v. Board of Education (1954): Segregation of students in public schools violates the Equal Protection Clause of the Fourteenth Amendment, because separate facilities are inherently unequal.

— Matrixx Initiatives, Inc. v. Siracusano (2011): A plaintiff may state a claim against a pharmaceutical company for securities fraud under the Securities Exchange Act of 1934 based on the company’s failure to disclose reports of adverse events even when the reports do not disclose a “statistically significant” number of such adverse events.

Required Reading:

SGEP (259–282) —

9.1 Weight-of-the-Evidence Arguments in the Presentation and Interpretation of Data

9.1.1 A Case Study in Data Interpretation: *Brown v. Board of Education* (1954)

9.1.2 A Case Study in Data Interpretation: *Matrixx Initiatives, Inc. v. Siracusano* (2011)

Popular Articles —

Head Case: Can Psychiatry Be a Science? Louis Menand (*New Yorker*, March 1, 2010)

Talking Back to Prozac, Frederick C. Crews (*New Yorker*, December 6, 2007)

Do We Really Know What Makes Us Healthy? Gary Taubes (*New York Times*, September 16, 2007)

The Plastic Panic, Jerome Groopman (*New Yorker*, May 31, 2010)

John Rock's Error, Malcolm Gladwell (*New Yorker*, March 10, 2000)

Suggested Reading:

9.1 Suggested Reading on Data Presentation and Interpretation

9.1.1 Appendix: *Brown v. Board of Education* (Decided: May 17, 1954)

9.1.2 Appendix: *Matrixx Initiatives, Inc. v. Siracusano* (Decided: March 22, 2011)

Film: *The Manhattan Project* (42 minutes); *The Town that Never Was* (16 minutes)

Week 9: (Mis)reporting of Data

** Exam 2 given on Thursday (covers the second four weeks) **

Required Reading:

SGEP (283–305) —

10.1 The Social Construction of Statistics

10.2 Adjustments for Groups Not Comparable On a Variable, Such As Age

Suggested Reading:

10.0.3 Appendix: *P. Lorillard Co. v. Federal Trade Commission* (Court of Appeals Fourth Circuit; Decided: December 29, 1950)

Film: *The Manchurian Candidate* (127 minutes)

Week 10: Inferring Causality

— inferring causality; Bradford–Hill criteria; historic medical conceptions of disease causality; medical error as the causative factor

Required Reading:

SGEP (307–331) —

11.1 Causuistry

11.2 The Bradford–Hill Criteria for Determining a Causal Connection

11.3 Some Historical Health and Medical Conceptions of Disease Causality

11.4 Medical Error as a (the) Causative Factor

Popular Articles —

The Truth Wears Off, Jonah Lehrer (*New Yorker*, December 13, 2010)

Lies, Damned Lies, and Medical Science, David H. Freedman (*The Atlantic*, November 2010)

Film: *Smoking and Cigarette Film Collection* (102 minutes)

Week 11: Simpson's Paradox; Meta-Analysis

— the ubiquity of Simpson's Paradox in the (mis)interpretation of data; when a relationship that appears to be present at an aggregated level disappears or reverses when disaggregated and viewed within levels

— meta-analysis and the controversies it engenders in childhood sexual abuse and other medically relevant research summarizations;

Required Reading:

SGEP (333-357) —

Popular Article —

Meta-Analysis at 25; Gene V. Glass, January 2000

Suggested Reading:

12.1 Suggested Reading on Simpson's Paradox

13.1 Suggested Reading on Meta-analysis

Film: *Sacco and Vanzetti* (82 minutes)

Week 12: Statistical Sleuthing and Explanation

— statistical sleuthing with formal models: Poisson clumping, Benford's law, survival analysis, Kaplan–Meier curves;

— *McCleskey v. Kemp* (1987): Despite statistical evidence of a profound racial disparity in application of the death penalty, such evidence is insufficient to invalidate defendant’s death sentence.

Required Reading:

SGEP (359–384) —

14.1 Sleuthing Interests and Basic Tools

14.2 Survival Analysis

14.3 Statistical Sleuthing and the Imposition of the Death Penalty:
McCleskey v. Kemp (1987)

Popular Articles —

The Treatment, Malcolm Gladwell (*New Yorker*, May 17, 2010)

The Ghost’s Vocabulary: How the Computer Listens for Shakespeare’s “Voiceprint,” Edward Dolnick (*The Atlantic*, October, 1991)

Suggested Reading:

14.1 Suggested Reading on Statistical Sleuthing

14.1.1 Appendix: U.S. Supreme Court, *McCleskey v. Kemp* (Decided: April 22, 1987): Majority Opinion and Dissent

Film: *A Cry in the Dark* (121 minutes)

Week 13: Background: Experimental Design and the Collection of Data

** Exam 3 given on Thursday (covers the third four weeks)

— difficulties with observational studies; the hormone replacement therapy controversy and related artifactual interpretative issues

Required Reading:

SGEP (387–419) —

15.1 Observational Studies: Interpretation

15.2 Observational Studies: Types

15.3 Observational Studies: Additional Cautions

15.4 Controlled Studies

15.5 Controlled Studies: Additional Sources of Bias

15.5.1 The special case of medical trials

15.5.2 An introductory oddity: The 1954 Salk polio vaccine trials

15.6 The Randomized Response Method

Popular Articles —

Influence of Funding Source on Outcome, Validity, and Reliability of Pharmaceutical Research, Report 10 of the Council on Scientific Affairs of the American Medical Association

Sponsorship, Authorship, and Accountability: International Committee of Medical Journal Editors (August, 2007)

Suggested Reading:

15.1 Suggested Reading on the Production of Data and Experimental Design

Film: *The Fog of War* (107 minutes)

Week 14: Ethical Considerations in Data Collection

— ethical considerations in data collection and analysis involving human experimentation; the Nazi Doctors' Trial and the Nuremberg Code; the Tuskegee syphilis study and the Belmont Report; the Declaration of Helsinki and the conduct of foreign clinical trials

Required Reading:

SGEP (421–447) —

16.1 The Nazi Doctors’ Trial and the Nuremberg Code

16.2 The National Research Act of 1974

16.3 The Declaration of Helsinki

Popular Articles —

Whose Body is it, Anyway? Atul Gawande (*New Yorker*, October 4, 1999)

Drug Companies & Doctors: A Story of Corruption, Marcia Angell (*New York Review of Books*, January 15, 2009)

Suggested Reading:

16.0.1 Appendix: The Belmont Report

Film: *Miss Evers’ Boys* (118 minutes)

Week 15: The Federal Rules of Evidence; Some Concluding Remarks

— the Federal Rules of Evidence and the court admissibility of expert witnesses and scientific data; the Daubert trilogy of Supreme Court decisions

— the importance of context and framing in the presentation of data; the work of Tversky and Kahneman, and the more recent points made by Gigerenzer and his colleagues (“Helping Doctors and Patients Make Sense of Health Statistics,” in the series sponsored by the Association for Psychological Science, Psychological Science in the Public Interest)

— *Daubert v. Merrell Dow Pharmaceuticals* (1993): The Federal Rules of Evidence govern the admission of scientific evidence in a trial held in federal court. They require the trial judge to act as a gatekeeper before admitting the evidence, determining that the evidence is scientifically valid and relevant to the case at hand.

Required Reading:

SGEP (449–492)–

17.1 Junk Science

17.2 The Consequences of *Daubert* and the Data Quality Act (of 2001)

Popular Articles–

Science and Society: The Interdependence of Science and Law, Stephen Breyer (*Science*, April 24, 1998)

Something Rotten At the Core of Science? David F. Horrobin (*Trends in Pharmacological Sciences*, February, 2001)

Is Science Different for Lawyers? David L. Faigman (*Science*, July 19, 2002)

Scientific Evidence and Public Policy, David Michaels (*American Journal of Public Health*, Supplement 1, 2005)

Doubt Is Their Product, David Michaels (*Scientific American*, June, 2005)

Suggested Reading:

17.1 Suggested Reading on the Admissibility of Expert Testimony, The Federal Rules of Evidence, and Related Topics

Film: *West of Memphis* (2 hours)

Other Films of Interest:

The Atomic Cafe (88 minutes)

Salem Witch Trials (50 minutes)

The Crucible (123 minutes)

A Paralyzing Fear (90 minutes)

The Grapes of Wrath (129 minutes)

Time of Fear (60 minutes)

Good Night, and Good Luck (92 minutes)

Triumph of the Will (122 minutes)

The Dust Bowl (4 hours)

Minority Report (146 minutes)

Wild River (110 minutes)

Tomorrow's Children (52 minutes)

Why We Fight (7 films by Frank Capra)

December 7th (104 minutes)

Paradise Lost Trilogy (3 films)

The Most Dangerous Man in America: Daniel Ellsberg and the Pentagon Papers (94 minutes)

Salt of the Earth (94 minutes)

I Was A Communist for the FBI (82 minutes)

The Day After Trinity (88 minutes)

A Face in the Crowd (125 minutes)

Howard Zinn: You Can't Be Neutral on a Moving Train (78 minutes)

Henry A. Wallace: An Uncommon Man (57 minutes)

Scottsboro: An American Tragedy (90 minutes)

To Kill a Mockingbird (130 minutes)