November, 2015

Experts, Statistics, Science & Bad Science

Curtis E.A. Karnow

RESOURCES: EXPERTS, STATISTICS, SCIENCE & BAD SCIENCE

Curtis Karnow, Superior Court (San Francisco)

This was originally assembled for presentations at the California Complex Judges' meeting and related Bench-Bar conference of November 12, 2015, and is augmented here (primarily with materials on statistics)

Evidence & experts

- M. Simons, CALIFORNIA EVIDENCE MANUAL Ch.4 (Experts)
- M. Simons, "Introducing Hearsay Through An Expert: Is the Backdoor Closing?," 20 ABTL Report 1 (Summer 2011)
- CEB, ACTION GUIDE: HANDLING EXPERT WITNESSES IN CALIFORNIA COURT
- Curtis Karnow, "Sargon and the Science of Reliable Experts," 22 ABTL REPORT 1 (Spring 2013), original at http://works.bepress.com/curtis_karnow/
- David L. Faigman & Edward J. Imwinkelried, "Wading into the Daubert Tide: Sargon Enterprises, Inc. v. University of Southern California," 64 HASTINGS L.J. 1665, 1691 (2013) (authors of the law review article relied on in *Sargon*)
- Jules Epstein, "Preferring the "Wise Man" to Science: The Failure of Courts and Non-Litigation Mechanisms to Demand Validity in Forensic Matching Testimony," 20 WIDENER L. REV. 81, 82 (2014) (criticizing use of unproven forensic techniques such as latent prints and handwriting analysis)

Reporting bias & related issues (peer reviews)

- Christie Aschwanden, "Science Isn't Broken: It's just a hell of a lot harder than we give it credit for," *FiveThirtyEightScience* (August 19, 2015), http://fivethirtyeight.com/features/science-isnt-broken/ (superb review of rationales behind retractions of papers, why p values can be misleading, why peer reviewed journals don't guarantee reliability, etc. Provides an interactive chart that illustrates how p values can be easily manipulated)
- Daniele Fanelli, "How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data," http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0005738
- B. Goldacre, BAD SCIENCE (2008); see also his site, http://www.badscience.net
- Richard Smith, "Peer Review: a flawed process at the heart of science and journals," 99:4 *J.R. Soc.Med*. 171 (April 2006), available at http://jrs.sagepub.com/content/99/4/178.long
- http://retractionwatch.com/2014/06/30/how-often-do-economists-commit-misconduct/ [the site *retractionwatch.com* has a very useful listing of retracted papers]
- http://www.physics.ohio-state.edu/~wilkins/science-fraud.html
- http://michaelnielsen.org/blog/three-myths-about-scientific-peer-review/
- http://www.wsj.com/articles/hank-campbell-the-corruption-of-peer-review-is-harming-scientific-credibility-1405290747
- David L. Faigman, "Anecdotal Forensics, Phrenology, and Other Abject Lessons from the History of Science," 59 HASTINGS L.J. 979 (2008)
- S. Goldbeck-Wood, "Evidence on peer reviews-scientific quality or smokescreen?," http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1114539/ ("nationality bias, language bias, specialty bias, and perhaps even gender bias, as well as the recognised [sic] bias toward the publication of positive results")

- "The Houses of Deceits: Science, Forensic Science, and Evidence and Introduction to Forensic Evidence," 35 LAND & WATER L. REV. 397, 400-01 (2000)
- Richard Wilson, "Technology and Society: Ensuring Sound Science in the Courts" (Harvard: 2003 draft), http://users.physics.harvard.edu/~wilson/publications/ppaper871.html#N_2_
- Richard van Noorden, "The Trouble with Retractions," 478 NATURE 26 (October 6, 2011)
- Benedict Carey, "Science, Now Under Scrutiny Itself," THE NEW YORK TIMES http://nyti.ms/1Gom501 (June 15, 2015)
- Ben Goldacre, "Scientists Are Hoarding Data And It's Ruining Medical Research: Major flaws in two massive trials of deworming pills show the importance of sharing data — which most scientists don't do,"(Jul. 22, 2015), http://www.buzzfeed.com/bengoldacre/deworming-trials
- http://www.sciencemag.org/content/348/6239/1100.2.full (retractions of scientific papers)
- http://www.nature.com/news/2011/111005/pdf/478026a.pdf (high increase in retractions of scientific papers)
- Christine Schmucker, "Extent of Non-Publication in Cohorts of Studies Approved by Research Ethics Committees or Included in Trial Registries," http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0114023 (questioning validity of systematic reviews because journal publications represent a biased selection of all studies conducted [dissemination bias])
- Fujian Song, et al., "Dissemination and publication of research findings: an updated review of related biases," *Health Technol. Assess*. (2010), available at http://www.researchgate.net/profile/Fujian_Song/publication/41561626_Dissemination_and_publication_of_research_findings_an_updated_review_of_related_biases/links/09e4150b49a57536fc000000.pdf ("Studies with significant or positive results were more likely to be published than those with non-significant or negative results …There was convincing evidence that outcome reporting bias exists and has an impact on the pooled summary in systematic reviews. … published studies tended to report a greater treatment effect than those from the grey literature. Exclusion of non-English language studies appeared to result in a high risk of bias in some areas")
- See generally, Curtis Karnow, "Cognitive Fallacies Reading List," for more references on biases and fallacies which interfere with people's logical and scientific reasoning http://works.bepress.com/curtis_karnow/11/

Metastudies & reproducible results:

- *Cochran systematic review*: http://www.cochrane.org/cochrane-reviews; http://bmg.cochrane.org/addressing-reporting-biases (Cochrane furthers transparency in research and publication, and use of metastudies)
- Importance of meta studies: http://community.cochrane.org/about-us/evidence-based-health-care/webliography/books/sysrev
- http://www.alltrials.net/
- http://boingboing.net/2014/05/15/half-of-all-clinical-trials-ne.html
- Ben Goldacre, "Listen carefully, I shall say this only once," *The Guardian* (October 25, 2008), available at http://www.badscience.net/2008/10/listen-carefully-i-shall-say-this-only-once/#more-823 (problems with issuing multiple reports of what is, in fact one study; contrast results of the 'one' study with metastudy results)

Fake science papers:

- Kerry Grens, "Fake Paper Exposes Failed Peer Review," *The Scientist* (October 6, 2013), available at http://www.the-scientist.com/?articles.view/articleNo/37798/title/Fake-Paper-Exposes-Failed-Peer-Review/
- http://pdos.csail.mit.edu/scigen/ [create your own fake paper in seconds]
- http://www.nature.com/news/publishers-withdraw-more-than-120-gibberish-papers-1.14763 (120 gibberish papers withdrawn)
- http://scigendetection.imag.fr/main.php (possible detection of fake papers)
- "Publishers withdraw more than 120 gibberish papers: Conference proceedings removed from subscription databases after scientist reveals that they were computer generated," NATURE NEWS & COMMENT (June 23, 2015)
- http://www.michaeleisen.org/blog/?p=1439 (story of getting a fake paper accepted by prestigious journal)

Science

- Robert Ehrlich: NINE CRAZY IDEAS IN SCIENCE (2002), http://press.princeton.edu/chapters/i7022.pdf
- Richard Feynman, THE PLEASURE OF FINDINGS THINGS OUT (2005)
- Karl Popper, CONJECTURES AND REFUTATIONS: THE GROWTH OF SCIENTIFIC KNOWLEDGE (1963-2002)

Charts and graphs – the good, the bad, the ugly, and the beautiful

- http://www.edwardtufte.com/tufte/index. Professor Tufte also discusses and illustrates *useful* graphical representations of data and statistics in his articles and books such as:
  - "Visual and Statistical Thinking: Displays of Evidence of Making Decisions," https://www.sfu.ca/cmns/courses/2012/801/1-Readings/Tufte%20Visual%20and%20Statistical%20Thinking.pdf
  - VISUAL DISPLAY OF QUANTITATIVE INFORMATION (1983)
  - ENVISIONING INFORMATION (1990)
  - VISUAL EXPLANATIONS (1997)
  - BEAUTIFUL EVIDENCE (2006)
- https://visualisingadvocacy.org/blog/disinformation-visualization-how-lie-datavis
- http://flowingdata.com/
- Manuel Lima, VISUAL COMPLEXITY: MAPPING PATTERNS OF INFORMATION (2011) (a lovely collection of short essays and many illustrations of complex systems and massive data sets)
- Wonderful graphical representations of statistical evidence: http://www.gapminder.org/

Public science literacy (or absence of)

- "Major Gaps Between the Public, Scientists on Key Issues," *Pew Research* (July 2015) http://www.pewinternet.org/interactives/public-scientists-opinion-gap/ ("Despite broadly similar views about the overall place of science in America, there are striking differences between the views of the public and those of the scientific community connected to the American Association for the Advancement of Science (AAAS) on a host of science-related issues, from whether genetically modified foods are safe to eat to whether the world's growing population will be a major problem")
- "Public's Knowledge of Science and Technology," *Pew Research* (April 2013)
- Gallup 2005, http://www.gallup.com/poll/16915/three-four-americans-believe-paranormal.aspx

- Public Policy Polling (2013) http://boingboing.net/2013/04/15/12-million-americans-believe-l.html (ESP: 41%; Haunted houses: 37%; Ghosts: 32%; Telepathy: 31%; Astrology: 25%; Lizard people control politics: 4% (12,556,562); +7% who are *not sure if lizard people are involved* (?!))
- "Science and Technology: Public Attitudes and Understanding," National Science Board (2004, most are 2001 results), http://www.nsf.gov/statistics/seind04/c7/c7s2.htm#c7s2l5:
    - Two-thirds (in 2001) do not have a firm grasp of what is meant by the scientific process
    - At least a quarter of population believes in astrology
    - Europeans were more likely to say that astrology is scientific than to say the same about economics
    - Europeans (46%) Americans (32%) agree that "some numbers are particularly lucky for some people"
    - At least half of U.S. believes in the existence of extrasensory perception (ESP)
    - 30% agreed that "some of the unidentified flying objects that have been reported are really space vehicles from other civilizations."
- Eula Biss, ON IMMUNITY: AN INOCULATION (2014) (why educated middle class Americans irrationally fear immunizations)

Statistics

- See, Christie Aschwanden, "Science Isn't Broken" under **Reporting bias & related issues** above
- Christie Aschwanden, "Statisticians Found One Thing They Can Agree On: It's Time To Stop Misusing P-Values," (March 7, 2016), http://fivethirtyeight.com/features/statisticians-found-one-thing-they-can-agree-on-its-time-to-stop-misusing-p-values/ (good discussion of what p values don't mean)
- Jordan Ellenberg, HOW NOT TO BE WRONG: THE POWER OF MATHEMATICAL THINKING (2015) (broad introduction to practical mathematical literacy, including materials on statistics)
- Edward Cheng, "Fighting Legal Innumeracy," 17 GREEN BAG 2D 271 (Spring 2014) (reasons and plea for better understanding of statistics in the legal profession)
- George Akerlof & Robert Schiller, PHISHING FOR PHOOLS: THE ECONOMICS OF MANIPULATION AND DECEPTION (2015) (an entirely accessible review by two Nobel Prize winners of ways in which numbers are used to fool the public, information asymmetry in the markets, and the role of regulation; not a detailed review of the misuse of statistics as such)
- Alex Reinhart, STATISTICS DONE WRONG: THE WOEFULLY COMPLETE GUIDE (2015)
- Sherry Seethaler, LIES, DAMNED LIES, AND STATISTICS (2009)
- C. Seife, PROOFINESS (2010) (subtitled, "The Dark Art of Mathematical Deception")
- G. Gigerenzer, RISK SAVVY (2014)
- C. Wheelan, NAKED STATISTICS (2013)
- Megan Higgs, "Do We Really Need the S-word?," AMERICAN SCIENTIST (Jan.-Feb. 2014), available at http://www.americanscientist.org/issues/pub/do-we-really-need-the-s-word/1 (noting problems misunderstanding 'significance' in connection with p values)
- Curtis Karnow, "Statistics In Law: Bad Inferences & Uncommon Sense" (2011) http://works.bepress.com/curtis_karnow/
- Gary Smith, STANDARD DEVIATIONS (2014)
- Timothy Urdan, STATISTICS IN PLAIN ENGLISH (3d ed. 2010)
- Ian Ayers, SUPER CRUNCHERS (statistics and other use of large numbers in a variety of disciplines, written for a general audience)
- John Phillips, Jr., HOW TO THINK ABOUT STATISTICS (1971)
- N.N. Taleb, THE BLACK SWAN: THE IMPACT OF THE HIGHLY IMPROBABLE (2010)

- Marcy M. Hallock, "The Numbers Game - The Use and Misuse of Statistics in Civil Rights Litigation," 23 Vill. L. Rev. 5 (1978)
- Regina Nuzzo, "Scientific method: statistical errors," *Nature,* February 12, 2014, http://www.nature.com/news/scientific-method-statistical-errors-1.14700 (weakness of just using p values)
- A.K. Dewdney, 200% OF NOTHING (1993) (subtitled "An Eye-Opening Tour Through the Twists and Turns of Math Abuse and Innumeracy")
- "Just Plain Data Analysis: Common Statistical Fallacies in Analyses of Social Indicator Data," also found here: http://lilt.ilstu.edu/jpda/interpreting/interpreting_the_numbers.htm
- See sample misleading statistics at http://www.econoclass.com/misleadingstats.html
- Statistical fallacies: http://lilt.ilstu.edu/jpda/interpreting/interpreting_the_numbers.htm
- Type I and type II errors in the context of the criminal justice system: http://intuitor.com/statistics/T1T2Errors.html
- D. Kaye et al., Reference Guide on Statistics, http://ftp.resource.org/courts.gov/fjc/sciam.0.stats.pdf
- "The Use -- and Misuse -- of Statistics: How and Why Numbers Are So Easily Manipulated," (April 2, 2008), located at Knowledge@Wharton
- Excellent series of videos on many subjects including probability and statistics: kahnacademy.org
- Statistical evidence of forgery: http://en.wikipedia.org/wiki/Howland_will_forgery_trial
- *Walking through the Bayesian theorem:*
    - http://www.vjs.org/spam/bayesian-analysis.html
    - http://www.kilty.com/class8.htm
    - Peter Sedlmeier & Gerd Gigerenzer, "Teaching Bayesian Reasoning in Less Than Two Hours," 130 *Journal of Experimental Psychology: General* (2001) 380-400.
- Freedman et al., STATISTICS (4th ed.2007); Kaye & Freedman, 'Reference Guide on Statistics', in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE (3d ed.2011) relied on by the Supreme Court in *Duran v. U.S. Bank Nat. Assn*., 59 Cal. 4th 1, 38, 43 (2014); see also earlier edition, D. Freeman, STATISTICS (3d ed. 1998). The 1061 page MANUAL, which is found at http://www.fjc.gov/public/pdf.nsf/lookup/SciMan3D01.pdf/$file/SciMan3D01.pdf, also includes:
    - "Reference Guide on Multiple Regression"
    - "Reference Guide on Survey Evidence"

A few cases

- Case in which the rejected statistical survey used the wrong population for sampling: *Citizens Financial Group, Inc. v. Citizens Nat. Bank of Evans City*, 383 F.3d 110, 120 (3d Cir. 2004)
- Case in which the trial court was reversed for having fallen into a statistical trap (a 'logical fallacy,' the appellate court called it): *Sylvia Darensburg v. Metropolitan Transportation Commission* (9th Cir., February 16, 2011, No. 09-15878)
- Court takes judicial notice of certain statistical facts: *Envtl. Law Found. v. Beech-Nut Nutrition Corp.,* 235 Cal. App. 4th 307, 325 n.7 (2015)
- Trial plan failures, including failure to present viable statistical plan, support denial of class certification: *Mies v. Sephora U.S.A., Inc.,* 234 Cal. App. 4th 967 (2015)
- "Statistical significance" discussed: *Lewis v. Ascension Par. Sch. Bd*., 806 F.3d 344, 362 n.24 (5th Cir. 2015); *Chen-Oster v. Goldman, Sachs & Co*., 114 F. Supp. 3d 110 & n.4 (S.D.N.Y. 2015)
- Problems with sample size: *E.E.O.C. v. Freeman*, 778 F.3d 463, 469 & n.1 (4th Cir. 2015)

- Use of regression analysis: *E.g., In re Se. Milk Antitrust Litig.,* 739 F.3d 262, 285 (6th Cir.) *cert. denied sub nom. Dean Foods Co. v. Food Lion, LLC*, 135 S. Ct. 676 (2014); *Werdebaugh v. Blue Diamond Growers*, 2014 WL 2191901 (N.D. Cal. May 23, 2014) (proving damages under UCL, FAL, and CLRA); *Kleen Products LLC v. Int'l Paper*, 306 F.R.D. 585, 602 (N.D. Ill. 2015)

∞

Statistics: On-line glossaries and basic introductions

http://www.stats.gla.ac.uk/steps/glossary/presenting_data.html#med
http://stats.oecd.org/glossary/
http://en.wikipedia.org/wiki/Glossary_of_probability_and_statistics
http://bobhall.tamu.edu/FiniteMath/Module8/Introduction.html
http://stattrek.com/Help/Glossary.aspx
http://statistics.berkeley.edu/~stark/SticiGui/Text/gloss.htm
http://www.statsoft.com/textbook/basic-statistics/
http://www.statpac.com/statistics-calculator/correlation-regression.htm

Calculators

http://easycalculation.com/statistics/statistics.php http://www.surveysystem.com/sscalc.htm
http://www.dimensionresearch.com/resources/calculators/conf_means.html
http://www.stat.tamu.edu/~jhardin/applets/
E.g. calculate p value:
http://www.ehow.com/how_5073193_calculate-p_values-t_tests.html

∞

A few notes on standard deviation:

Assume the SD for IQ (intelligence) is 15; then 2 SD = 30. So 95% (about 2 SD) people have an IQ between 70 (100 = mean, ± 15 x 2) and 130 (= 100 + 2 x 15)
If what we're measuring is less variable, say with a SD=5, then the 95% [or 2 SD range] would be 100 ± 10, i.e., 90-110

A few notes on the value of p:

High p is bad, low p is good.
P = 0.05 is generally borderline acceptable error, statistically significant for many uses.
P < 0.01 commonly considered statistically significant.
P ≤ 0.005 or p ≤ 0.001 means the findings are highly statistically significant.
The significance level of a statistical hypothesis test is a fixed probability of wrongly rejecting the null hypothesis,[1] if it is in fact true.
Often, the significance level is chosen to be 0.05 (or equivalently, 5%).

[1] The null hypothesis is the opposite of the claim being tested. I.e. if the hypothesis is that Drug X reduces cancer, then the null hypothesis is that Drug X does not reduce cancer.

These phrases all these mean roughly the same thing:

The finding is significant at the .05 level.
The confidence level is 95 percent.
There is a 95 percent certainty that the result (testing the null) is not due to chance.
There is a 1 in 20 chance of obtaining this result by chance.
The *p*-value is .05.

A classic error in understanding "margin of error":

Assume we have poll that ranks two candidates: Laurel comes in at 52%, and Hardy at 48%. We told there is a "margin of error" of ± 2%.
Does this mean that in fact the two candidates are too close to call? Because both candidates could be, with a margin of error at ± 2%, at 50%?
No. That intuition is misleading.
Here's how we interpret the results. Let's assume we have the classic two standard deviation results (that's where the 95% chance language below comes from):
Laurel has a 95% chance that the REAL number (in the population) is somewhere between 50-54%
Hardy has a 95% chance that the REAL number (in the population) is somewhere between 46-50%.
So: Laurel is very, very likely the winner in the real world/population.

A note on the bizarre Newcomb-Benford Law:

As we study randomness and seek to develop our intuitions, it is worth repeating that our sense of what is, and is not, random, is often at odds with reality.
One of the most bizarre examples of this is the Newcomb-Bedford law, which notes that when we look at the first significant figures of a number (i.e. '3' in 385 or '7' in 785,945) smaller numbers predominate over larger numbers in a wide variety of real world circumstances. For example, if we look at the leading number of bank accounts, stock prices, numbers on tax returns, the areas and population of countries, and the starting page numbers of papers from a bibliography, and many other numbers, we will see that the smaller the number, the more frequently it occurs. "1" appears almost 1/3 of the time, although we might think that, randomly, it would appear about 1/9 of the time. "1" appears most frequently, "2" next in frequency, and so on. In short, apparently random numbers may actually be evidence of artificiality. As a result, statistical studies of check amounts, numbers in tax returns, and other figures can detect artificial numbers-- that is, fraud.
See, e.g., http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0010541
http://www.lynceangroup.net/BenfordLynceanPresentation.pdf
The technique has been used to detect vote fraud. Peter Klimek, et al., "Statistical detection of systematic election irregularities," *Proc Natl Acad Sci U S A*. 109(41): 16469–16473 (2012), http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3478593/. Others have suggested Benford's law is not useful in this way, Joseph Deckert, et al., "Benford's Law and the Detection of Election Fraud," *Political Analysis* 19 (3):245-268 (2011), http://pan.oxfordjournals.org/content/19/3/245.abstract. *Compare*, e.g., Walter R. Mebane Jr., "Comment on "Benford's Law and the Detection of Election Fraud," *Political Analysis* (2011) 19 (3):269-272 (2011), http://pan.oxfordjournals.org/content/19/3/269.abstract (severe criticism of Deckert article)

∞

Deciphering notations and terms:

Confidence level of 95% corresponds to a significance level (p) of 5%; a confidence level of 99% corresponds to a significance level of 1%.

± = plus or minus

Σ = 'sum' and indicates one is to add up (or sum) the indicated numbers, as follows:

$$\sum_{i=1}^{n}$$

This indicates that one "sums" (or adds up) the numbers starting with 1 (because i indicates the number to start with, which here =1) and go the last number, which is n.

So *if we're told n=3*, then this:

$$\sum_{i=1}^{N}$$

is the equivalent to: 1+2+3, i.e., 6.

μ = population mean (*pronounced "mu"*)
which is in turn calculated with this formula:
$(\Sigma x_i) / n$
This says: the sum of (Σ) all the numbers x_1 to x_2, x_3 through the end, then divide by the number of x's ('n'). So if all our x's (sample data) are 1, 4, 7, and 8, then we sum them (= 20) and we divide by the number of samples we have, 4, so the result is: μ=5.

σ = Population standard deviation = $\text{sqrt} [\Sigma (X_i - \mu)^2 / N]$

σ squared (or σ^2) = variance

$\bar{x}$ = sample mean
which is calculated this way:
$(\Sigma x_i) / n$

P = probability
P(A) = probability of event A
P(A | B) = probability of A given B has occurred
∩ = intersection = "and"
U = union = "or"
P(A or B) = P(A U B) = probability of A or B occurring
P(A and B) = P(A ∩ B) = probability of A and B both occurring

For further explanations of symbols and terms, see e.g.
www.stat.tamu.edu/~julie/302/handouts/**symbols**.doc