Defensive Statistics

John Armstrong, Spring 2018

Many of the statistical assertions in the media are misleading and/or wrong. This applies to medical and diet-related articles, as well as to much election polling, and many other topics. (Most reporters do not understand statistics, or if they do understand, aren't motivated or allowed to explain the potential errors involved.)

A particular example: all diagnostic medical tests exhibit false-positive and false-negative rates. Home pregnancy tests are a case in point. These inevitable features do not invalidate testing, but they show the need for informed caution. Ideally, participants in this seminar will acquire a set of tools with which to approach the many statistical assertions in the media.

Possible Topics

-Choose a chapter from **"How to Lie with Statistics"**, summarize, and give contemporary examples.

-Describe the various **ways in which graphs of statistical data can mislead**, using examples from contemporary media.

-What are the different **kinds of errors** that **poling** is subject to? What are **'frame errors' in polling?** Provide examples.

-The topic of statistics is inseparable from that of **measurement errors**. Describe the ways in which errors can accumulate, e.g. when determining the difference between two large, measured quantities, or when taking the the ratio of two measured quantities.

-**Diagnostic testing for disease** is widespread, but the role of **test-error-rates** is not widely understood. Present on this topic.

-The field of **experimental psychology** has recently been shaken by a **'reproducibility crisis'**; Lack of reproducibility is often related to weak statistical methodology. Report on this matter.

-The concept of **randomness** is crucial to many statistical projects; randomness includes the occurrence of **clusters**. Report on the connection between randomness and clusters. This is often relevant to epidemiological studies.

-There is a long-running controversy about the fundamental meaning of statistics. This controversy pits the **'frequentist' followers of R.A. Fisher** and others

against those of the Rev. Thomas **Bayes** (1701-1761). Report on this controversy.

-What are the statistical conditions under which experimental data can justify drawing the conclusion that a causal relationship has been demonstrated?

-Suggest your own topic in consultation with the moderator.

Resources

How to Lie with Statistics, Darryl Huff (available in paperback)
Statistics for Dummies, Deborah H. Rumsey (paperback)
Many web articles; clips from magazines & daily newspapers…

Ionnidis' papers

"Why Most Published Research Findings are False."
http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0020124

"Economics isn't a bogus science — we just don't use it correctly."
http://www.latimes.com/opinion/op-ed/la-oe-ioannidis-economics-is-a-science-20171114-story.html