Lectures on Foundations of Statistics and Inference

P.B. Stark www.stat.berkeley.edu/~stark

This document: www.github.com/pbstark/Kavli18/kavliStat18.pdf

Last edited 9 July 2018

Reading for the week, and other references

1. Boring, A., K. Ottoboni, and P.B. Stark, 2016. Student evaluations of teaching (mostly) do not measure teaching effectiveness, *ScienceOpen Research*, DOI: 10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1 https://www.scienceopen.com/document?vid=818d8ec0-5908-47d8-86b4-5dc38f04b23e

2. Freedman, D.A., 1995. Some issues in the foundations of statistics, *Foundations of Science*, *1*, 19–39. https://doi.org/10.1007/BF00208723

3. Freedman, D.A., 2009. *Statistical Models: Theory and Practice*, 2nd edition, Cambridge University Press.

4. Freedman, D.A., R. Pisani, and R. Purves, 2007. *Statistics, 4th edition*, W.W. Norton, New York.

5. Klemes, V., 1989. The Improbable Probabilities of Extreme Floods and Droughts, in O. Starosolsky and O.M. Meldev (eds), *Hydrology and Disasters*, James and James, London, 43–51. https://www.itia.ntua.gr/en/getfile/1107/1/documents/1997_ImprobProbabilities_OCR.pdf

6. LeCam, L., 1977. Note on metastatistics or 'An essay toward stating a problem in the doctrine of chances,' *Synthese*, *36*, 133-160.

7. Morabia, A., 2006. Pierre-Charles-Alexandre Louis and the evaluation of bloodletting, *J. Roy. Soc. Medicine*, *99*, 158–160. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1383766/pdf/0158.pdf

8. Mulargia, F., P.B. Stark, and R.J. Geller, 2017. Why is probabilistic seismic hazard analysis (PSHA) still used?

9. Stark, P.B., and D.A. Freedman, 2003. What is the Chance of an Earthquake?, in *Earthquake Science and Seismic Risk Reduction*, F. Mulargia and R.J. Geller, eds., NATO Science Series IV: Earth and Environmental Sciences, v. 32, Kluwer, Dordrecht, The Netherlands, 201-213. https://www.stat.berkeley.edu/~stark/Preprints/611.pdf

10. Stark, P.B., 1997. SticiGui Various chapters assigned below.

11. Stark, P.B., 2016a. Pay no attention to the model behind the curtain

12. Stark, P.B., 2016b. The value of P-values, *The American Statistician*, *70*, DOI:10.1080/00031305.2016.1154108

13. Stark, P.B., 2017. Mathematical Foundations, Inequalities, Introduction to permutation tests, Rabbits and Cargo-Cult Statistics, Generating pseudo-random samples and permutations

14. Stark, P.B., and A. Saltelli, 2018. Cargo-cult Statistics and Scientific Crisis, *Significance*, *15*(4), 40–43. Preprint: https://www.significancemagazine.com/2-uncategorised/593-cargo-cult-statistics-and-scientific-crisis

15. Urban, M.C., 2015. Accelerating extinction risk from climate change, *Science*, *348*, Issue 6234, 571–573, DOI: 10.1126/science.aaa4984, http://science.sciencemag.org/content/348/6234/571.full

---

Lecture 1: Introduction to Inference.

The method of comparison, experiments and observational studies, Snow's investigation of the cause of cholera, Pierre Louis and bloodletting, confounding, randomization, controls, blinding, the Neyman model for causal inference, the Lady Tasting Tea experiment, Fisher's Exact Test, response schedules.

Assignment.

1. Read:
   - SticiGui Counting
   - SticiGui Theories of Probability
   - SticiGui Naive set theory
   - SticiGui Logic
   - SticiGui Axioms of Probability
   - SticiGui Experiments
   - Morabia (2006)
2. Work the self-test problems in the chapters of SticiGui assigned above.
3. **To hand in:** Look at the data Morabia transcribed from P.C.A. Louis on bloodletting for pneumonia. Is Louis's work an observational study or an experiment? Do you think it amounts to a "natural experiment," like Snow's work on Cholera? Why or why not? Give two scientific questions (*statistical hypotheses*) those data might address. What do you think the most important confounding factors would be, for those two hypotheses? What would be the most natural "as-if" randomization to use in analyzing the data to address the hypotheses you formulated, if you were to consider the data to be a natural experiment? What are the controls? Is the experiment blind? Double-blind? Explain how you might use the Neyman model to analyze the data. Is the randomization assumption reasonable? Why or why not? Is the non-interference assumption reasonable? Why or why not?

Lecture 2: Probability Models

Equally likely outcomes, the frequency theory, the subjective theory, probability as metaphor, probability models, Kolmogorov's axioms, consequences of the axioms, partitions, conditional probability, the multiplication rule, Bayes' rule, the law of total probability, useful probability inequalities. Box models, sampling distributions, the Law of Large Numbers, the expected value, common distributions arising from 0-1 boxes. Parameters, statistics, and estimation. Ontology of probability in applications. Bayesian and frequentist uncertainties; credible intervals and confidence intervals. Probability models in practice: earthquake probabilities, climate change and extinctions, birds and wind turbines.

Assignment

1. Read:
   - SticiGui Let's Make a Deal
   - SticiGui Probability Meets Data
   - SticiGui Random Variables
   - SticiGui Expectation
   - Mathematical Foundations
   - Inequalities
   - Freedman (1995)
   - Rabbits and Cargo-Cult Statistics
   - Klemes (1989)
   - LeCam (1977)
   - Stark and Freedman (2003)
   - Mulargia et al. (2017)
   - Urban (2015), including the supplementary materials
2. Work the self-test problems in the chapters of SticiGui assigned for this lecture.
3. **To hand in:** Urban estimates that 7.9% of species will become extinct as a result of climate change, with a 95% confidence interval of [6.2%, 9.8%]. Explain how Urban calculates the estimate and uncertainty. Is his work based on a random sample of species? Of geography? Explain his statistical model. Identify where in his work he assumes that associations are response schedules. (Hint: one of his references is McDonald, K.A., and J.H. Brown, 1992. Using montane mammals to model extinctions due to global change. *Conserv. Biol.*, *6*, 409–415. doi:10.1046/j.1523-1739.1992.06030409.x) For each such assumption, say whether you think it is reasonable, and explain why. Some aspects of the paper are Bayesian. What are the priors? What do the uncertainties mean? Is the interval [6.2%, 9.8%] a confidence interval? List 5 sources of uncertainty that he

omits. Do you think those sources are smaller or larger than those he includes? Explain.

Lecture 3: Tests

Statistical hypotheses, null and alternative hypotheses, hypothesis tests, families of tests, P-values, the meaning of P-values, abuses of P-values, group invariances of the distribution under the null hypothesis, conditional hypothesis tests, test statistics, Application: gender bias in teaching evaluations.

Assignment

1. Read:
   - Stark, 2016b.
   - Stark and Saltelli, 2018.
   - Introduction to permutation tests
   - Generating pseudo-random samples and permutations
   - Boring, Ottoboni, and Stark, 2016.
2. **To hand in:** Implement permutation tests (in R or Python) for the hypotheses about bloodletting you formulated for the first assignment, using the data Morabia transcribed from Louis. Discuss numerical issues in implementing the permutation tests. What pseudo-random number generator did you use? How did you set the seed? How did you choose the number of random permutations to perform? (Is there a principle you might use to decide the number? If so, what?) What test statistic(s) did you use? Why did you pick those? How to they connect to the relevant alternative hypotheses? Do you think the results would be qualitatively different if you had used a different test statistic? Discuss your findings, including any weaknesses in your framing of the problem (the statistical assumptions) and in the numerical analysis.