December 13, 2021

Rutgers University–New Brunswick Statistics Department

16:960:690
Seminar on the History of Probability and Statistics
Fall 2022

Instructor: Glenn Shafer, `www.glennshafer.com`, gshafer@rutgers.edu

Time: Tuesdays, 12:10 pm – 3:10 pm

Place: Busch Campus

Deliverables: Class discussion (40%), presentation of an article (10%), presentation of term paper (10%), term paper (40%)

Academic Integrity: Students are expected to adhere to Rutgers' academic integrity policy, described at `http://academicintegrity.rutgers.edu/`.

Purchase Recommended: *The History of Statistics: The Measurement of Uncertainty before 1900*, by Stephen M. Stigler. Harvard, 1986.

Why study the history of our subject? By tracing the path of discovery, we may gain a better appreciation of the logic underlying our methods—a better understanding of what is essential. We may also discover paths not taken, forks in the road where discovery and growth were diverted by circumstances and considerations that are no longer compelling. And we may be less prone to fruitless recycling of old positions and arguments.

How will the class be conducted?

- By the end of the day each Monday, please e-mail the instructor a question about the week's reading. You will be asked to discuss the question in class, explaining what you understand and what you would like to understand more fully. As much as half of each class will be devoted to discussing these questions.
- By September 27, please choose an article on the history of probability and statistics (perhaps but not necessarily from the articles listed in the following schedule) to present in class.
- The instructor will provide an extensive list of books, articles, and internet resources on the history of probability and statistics, to help students think about their choice of a topic for their term paper. Each student should agree with the instructor on a topic by October 18. Students will briefly present their term papers in the last class meeting (December 13), and the written paper (perhaps twenty pages) will be due December 23.

Tentative schedule

Please come to class prepared to help discuss the primary reading. In general, the secondary reading is meant for students who want to delve more deeply into the particular topic. But you may base the question you ask on either the primary reading or on the secondary reading.

This schedule may change to reflect students' interests, the instructor's own changing interests, and the availability of guest speakers.

1. September 6. Beginnings of the Calculus of Chances. The 13th century poem *De vetula* is the oldest known document in which chances are correctly calculated for a throw of three dice. It gives us a glimpse into the antiquity of calculation in games of chance and the equal antiquity of the interplay between belief (or betting) in the interpretation of the numbers. The numbers were not called probabilities.

Primary Reading

- Marie-France Bru and Bernard Bru (2018): Dice Games, *Statistical Science* 33(2):285–297.
- Glenn Shafer (2019): Pascal's and Huygens's game-theoretic foundations for probability, *Sartoniana* 32:117–145.

Secondary Reading

- Ian Hacking (2006): *The Emergence of Probability*, Second Edition, Introduction and Chapters 1–2. 40 pp.
- James Franklin (2001): *The Science of Conjecture: Evidence and Probability Before Pascal*, Chapter 11. Dice. 32 pp.
- David R. Bellhouse (2000): *De vetula*: a medieval manuscript containing probability calculations, *International Statistical Review* 68(2):123–136.

2. September 13. Beginnings of Bernoullian (a.k.a. Frequentist or Non-Bayesian) Inference. It began with Jacob Bernoulli's theorem (now called the law of large numbers for the binomial) in his posthumously published *Ars conjectandi* (1713) and Abraham De Moivre's theorem (now called the central limit theorem for the binomial) in 1733.

Primary Reading

- Jacob Bernoulli (1713): *Ars conjectandi*, Chapters IV and V of Part IV. On pp. 326–339 of Edith Sylla's 2006 translation, entitled *The Art of Conjecturing*.
- Stephen M. Stigler (1986): *The History of Statistics*, pp. 62–98. (Chapter 2. Probabilists and the measurement of uncertainty.)

Secondary Reading

- James Franklin (2001): *The Science of Conjecture: Evidence and Probability Before Pascal*, Chapter 10. Aleatory contracts: Insurance, annuities, and bets. 31 pp.
- Stephen E. Fienberg (1992): A brief history of statistics in three and one-half chapters: A review essay *Statistical Science* 7(2):208–225.

3. September 20. Beginnings of Non-Additive Numerical Probability. In *Ars conjectandi*, Bernoulli proposed that probabilities for repeatable events be found from frequencies. These probabilities would be additive, as are chances in games of chance. But for non-repeatable events, he could only construct arguments, and this lead him to possibly non-additive probabilities. Others had the same idea.

Primary Reading

- Jacob Bernoulli (1713): *Ars conjectandi*, Chapters I–III of Part IV. On pp. 315–326 of Edith Sylla's 2006 translation, entitled *The Art of Conjecturing*.
- Glenn Shafer (1986): The Combination of Evidence, *International Journal of Intelligent Systems* 1:155–179.

Secondary Reading

- Glenn Shafer (1978): Non-additive probabilities in the work of Bernoulli and Lambert, *Archive for History of Exact Sciences* 19:309–370.

4. September 27. Beginnings of Bayesian Inference. Pierre Simon Laplace invented inverse probability (Bayesian inference) in the 1770s, perhaps unaware of its previous invention by Thomas Bayes, or perhaps after having glimpsed Bayes's idea via Condorcet. Whereas Bayes's paper had received little notice until the 20th century, Laplace's version of the idea became an accepted part of the probability calculus.

Primary Reading

- Thomas Bayes (1763): An essay towards solving a problem in the Doctrine of Chances. *Philosophical transactions of the Royal Society of London* 53:370–418.
- Stephen M. Stigler (1986): Laplace's 1774 Memoir on Inverse Probability. *Statistical Science*, 1(3):359–363.

Secondary Reading

- Stephen M. Stigler (1986): *The History of Statistics*, pp. 99–138. (Chapter 3. Inverse Probability.)

- Glenn Shafer (1982): Bayes's two arguments for the rule of conditioning. *Annals of Statistics*, 10:1075–1089.

5. October 4. Laplace's Large-Sample Theory. In 1810, Laplace discovered how to approximate the distribution of a weighted average of a large number of random variables. This justified probability intervals based on the method of least squares, which can be given a Bernoullian interpretation (central limit theorem) or a Bayesian interpretation (Bernstein-von Mises theorem). In the Bernoullian interpretation, the distributions of the individual variables do not matter and need not be known. In the Bayesian interpretation, the prior probabilities are also irrelevant.

Primary Reading

- Stephen M. Stigler (1986): *The History of Statistics*, pp. 139–158. (Chapter 4. The Gauss–Laplace Synthesis.)

Secondary Reading

- Prakash Gorroochurn (2016): *Classic Topics on the History of Modern Mathematical Statistics: From Laplace to More Recent Times*, pp. 159–184.

6. October 11. Gauss's Theory of Estimation. Although Anne-Marie Legendre was the first mathematician to publish the method of least squares, the much younger Carl Friedrich Gauss claimed to have discovered it first, and he certainly did more with it. In 1823, he showed that least squares produced the best unbiased linear estimates of unknown constants; because Jerzy Neyman thought Gauss had been the first to prove this, it is now called the Gauss-Markov theorem. In general, Gauss and his followers were more concerned with accuracy of small-sample estimation than the French and English authors who followed Laplace.

Primary Reading

- Stephen M. Stigler (1981): Gauss and the Invention of Least Squares. *Annals of Statistics*, 9(3):465–474.

Secondary Reading

- Oscar Sheynin (2014): Elementary Exposition of Gauss' Final Justification of Least Squares. Slaski Przeglad Statystyczny 18(12):39–47.

7. October 18. Galton's Regression to Mediocrity. With an eye to improving the English race, Francis Galton undertook to turn his cousin Charles Darwin's theory of evolution into mathematics. The result was the theory of correlation and regression.

Primary Reading

- Stephen M. Stigler (2017): Regression. (Chapter 5, pp. 107-136 of *The Seven Pillars of Statistical Wisdom*, Harvard.)
- Francis Galton (1907): Probability, the Foundation of Eugenics, *Popular Science Monthly* 71:165–78.

Secondary Reading

- Stephen M. Stigler (1986): *The History of Statistics*, pp. 265–299. (Chapter 8. The English Breakthrough: Galton.)

8. October 25. Pearson and Yule Make Least Squares a Theory of Causality. Karl Pearson subscribed to Laplace's inverse probability. But his successors Ronald A. Fisher (1890–1962) and Jerzy Neyman (1894–1981), who put mathematical statistics into its present-day form, rejected inverse probability and its subjective philosophy in favor of methods and philosophy now called Bernoullian or frequentist.

Primary Reading

- Stephen M. Stigler (1986): *The History of Statistics*, pp. 326–361. (Chapter 10. Pearson and Yule.)
- John Aldrich (1995): Correlations genuine and spurious in Pearson and Yule, *Statistical Science* 10(4):364–376.

Secondary Reading

- George Udny Yule (1899): An investigation into the causes of changes in pauperism in England, chiefly during the last two intercensal decades (Part I.), *Journal of the Royal Statistical Society* 62(2):249–295.

9. November 1. Fisher's Small-Sample Theory. In contrast with the economists and other researchers who misapplied and discredited Laplace's large-sample theory in the 19th century, Karl Pearson and his biometricians often obtained their data by sampling. Because the sampling was often reasonably random, Fisher could argue that parametric inference for a small sample was reasonable so long a goodness-of-fit test did not show the sample was atypical of the assumed parametric model.

Primary Reading

- Ronald A. Fisher (1922): On the mathematical foundations of theoretical statistics, *Philosophical Transactions of the Royal Society of London. Series A*, 222:309–368.
- Stephen M. Stigler (2005): Fisher in 1921, *Statistical Science* 20:32–49.

Secondary Reading

- Stephen M. Stigler (2007): The Epic Story of Maximum Likelihood, *Statistical Science* 22(4):598–620.
- John Aldrich (2005): Fisher and Regression, *Statistical Science* 20(4):401–417.

10. November 8. Medical Statistics. The James Lind Library has pulled together much of the work on the history of medical statistics. Information about advances in this field are not usually well integrated into histories of mathematical statistics.

Primary Reading

- Nigel Paneth, Ezra Susser, and Mervyn Susser (2002): Origins and early development of the case-control study, Social and Preventive Medicine Soz Präventivmed 47:282–288, 359–365.
- Ulrich Tröhler (2011): The introduction of numerical methods to assess the effects of medical interventions during the 18th century: a brief history, *Journal of the Royal Society of Medicine* 104(11):465–474.

11. November 15. Revival of Bayes. As Sharon B. McGrayne's 2011 book documented, Laplace's inverse probability had many practitioners even as its opponents R. A. Fisher and Jerzy Neyman dominated mathematical statistics in the mid-20th century. By the 1960s, a subjectivist interpretation of probability newly invigorated by decision theory restored Bayes rule to a central place in mathematical statistics.

Primary Reading

- Stephen E. Fienberg (2006): When Did Bayesian Inference Become Bayesian?. *Bayesian Analysis* 1:1–40.
- Sharon B. McGrayne (2011): Jerome Cornfield, Lung Cancer, and Heart Attacks, Chapter 8 (pp. 108–118) of *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy*.

Secondary Reading

- Jerzy Neyman (1977): Frequentist probability and frequentist statistics. *Synthese*, 36:97–131.
- Erich L. Lehmann (2011): *Fisher, Neyman, and the Creation of Classical Statistics*. 115 pp.

12. November 29. Invention of Measure-Theoretic Probability. In the 1920s and 1930s, while the British were leading the development of statistics, the French and Russians developed a new foundation for probability and stochastic processes. Two celebrated milestones were Andrei Kolmogorov's *Grundbegriffe der Wahrscheinlichkeitsrechnung*, published in 1933, and Joseph L. Doob's *Stochastic Processes*, published in 1953. Jerzy Neyman was a major proponent of the thesis that mathematical statistics needed the new foundation for probability.

Primary Reading

- Glenn Shafer and Vladimir Vovk (2013): The origins and legacy of Kolmogorov's *Grundbegriffe*. 107 pp.
- Jerzy Neyman (1960): Indeterminism in science and new demands on statisticians. *Journal of the American Statistical Association*, 55:625–639.

Secondary Reading

- Jan von Plato (1994): *Creating Modern Probability*, Chapter 7. 46 pp.
- Andrei Kolmogorov (1933, translation 1956): *Foundations of the Theory of Probability.* 96 pp.

13. December 6. Invention of Game-Theoretic Probability. In his 1939 dissertation, in French, Ville used a betting game to cast light on the denumerable probabilities that Émile Borel had introduced in 1909 and that evolved into the measure-theoretic framework for probability used by mathematicians today. Ville's martingales contributed to two other complementary perspectives on mathematical probability: the understanding of randomness in terms of complexity, and the game-theoretic foundation for probability and mathematical statistics.

Primary Reading

- Glenn Shafer (2022): Did Jean Ville Invent Martingales? 16 pp. (Chapter 5 of *The Splendors and Miseries of Martingales: Their History from the Casino to Mathematics*, edited by Laurent Mazliak and Glenn Shafer, Birkäuser.)
- Glenn Shafer (2021): Testing by betting: a strategy for statistical and scientific communication, with discussion and response, *Journal of the Royal Statistics Society, Series A* 184(2):407-478, 2021.

Secondary Reading

- Laurent Bienvenu, Glenn Shafer, and Alexander Shen (2022): Martingales in the Study of Randomness. 38 pp. (Chapter 11 of *The Splendors and Miseries of Martingales: Their History from the Casino to Mathematics*, edited by Laurent Mazliak and Glenn Shafer, Birkäuser.)

14. December 13. Each student talks about their term paper.