Course Introduction

Aubrey Clayton

May 29, 2023

See lecture here.

1 History of Probability Theory

Mathematical notions of probability are thought to have originated from games and gambling. Dates back to letters exchanged between Pascal and Fermat trying to figure out the fair odds of throwing dice. Kolmogorov was the first to formalize probability (1933), which is a subset of measure theory.

He introduced probability by talking about the linguistic origins. He's arguing that it's pretty baked into our language (eg, when casually talking about odds of something happening). He's also pointing to the idea that the term *probability* is a bit ambiguous as to whether we're referring either to (i) a formal 'law' or way of making decisions or setting beliefs, or (ii) a way of describing frequencies of events unfolding over many observations.

2 History of Statistics

The term *statistics* originally referred to "of the state", or estimations that were of interest to the state. Understood that these estimations were subject to error such as measurement error or sampling bias. Another difficult question was how to combine different observations of things like astronomical data. Developmed most prominently by Laplace. These guys borrowed from probability theory to more accurately describe measurement errors—the language here is very much frequentist. A nice way of summing up statistics vs. probabilities is one is trying to quantify measurement frequencies and the other is trying to quantify frequencies of events.

Pearson and Neyman were working in parallel to and or competition with Fisher to develop a way of assesesing measurement error, and together developed the so-called orthodox statistics.

3 General Discussion

Interpretations of probability.

He quickly introduces Bayes, which is represented on both sides of his little probability / statistics chart.

He makes the interesting note that Kolmogorov's work doesn't really adjudicate between frequentist / Bayesian. It's just a set of rules for how probability is manipulated once a sample space is defined—it doesn't really give you a way to think about what probability really is and how to conceptually map it on to the real world.

When the concept of probability got applied beyond games of chance (some speculation), it really takes on this meaning of *degrees of belief.* A concern then is that degrees of belief are inherently subjective.

Scientific practice.

Hm okay now he's saying that Kolmogorov's axioms really only make sense when you're considering frequencies and ratios—they wouldn't really make sense if you're talking about degrees of belief. Jaynes reconciles this tension between probability and statistics by formalizing probability as an extension of logic, but in cases where we're reasoning with incomplete knowledge (ie, not deductions). Jaynes derives the Kolmogorov axioms from more fundamental logical reasoning. In frequentist statistics, conditional probability is kind of a weird offshoot, and Bayes rule comes from that. But when reasoning from Aristotelian logic, Bayes theorem comes very centrally.

Jaynes launches a huge critique of the Fisher frequentist methods. Frequentist logic starts from some null hypothesis, constructing a hypothetical distribution of summary statistics conditional on that hypothesis, and contrasts the observed data against that. Jaynes argues we need to start with the data as a given, and talk about the probability of the hypothesis given the observed data. Ie, instead of going hypothesis to data, we're going data to hypothesis.

Concludes by foreshadowing that Jaynes uses his Aristotelian logic and Bayes rule to dismantle the work of the last 200 years of orthodox statistics, and apparently Jaynes makes personal insults against Fisher throughout the book. Can't wait!