U.S. FOOD & DRUG ADMINISTRATION

INTRO TO BAYESIAN & FREQUENTIST STATISTICS

Jennifer L. Kirk

January 24, 2024

WCBP 2024

DISCLAIMER

This presentation is an informal communication and represents my own best judgement. My comments do not bind or obligate FDA.

STATISTICS

- Focused on drawing evidence or conclusions from data: information subject to uncertainty
- Much of statistics is underpinned by mathematics: probability, decision theory, etc.
- Bayesian and frequentist statistics are built on two different mathematical paradigms:
  - Different assumptions lead to different statistical methods for drawing conclusions from data

BAYESIAN & FREQUENTIST

- Frequentist
  - Probability: the long-run *frequency* at which an event occurs
  - Population parameters are fixed (i.e., take only a single value)
  - Frequentist methods designed to have specific properties in the long-run
    - Can be hard to design new methods, as no obvious way to design them
- Bayesian
  - Probability: the belief that an event will occur
  - Population parameters have distributions
  - Bayesian methods use Bayes' rule and shown to have specific properties
    - Easy to design new methods, but computationally intensive

BAYES' RULE (OR THEOREM)

$$p(A|B) = \frac{p(B|A)\, p(A)}{p(B)}$$

BAYESIAN STATS

$$p(\text{parameter} \mid \text{data}) = \frac{p(\text{data} \mid \text{parameter})\, p(\text{param.})}{p(\text{data})}$$

BAYESIAN STATS

posterior distribution
best guess after observed data

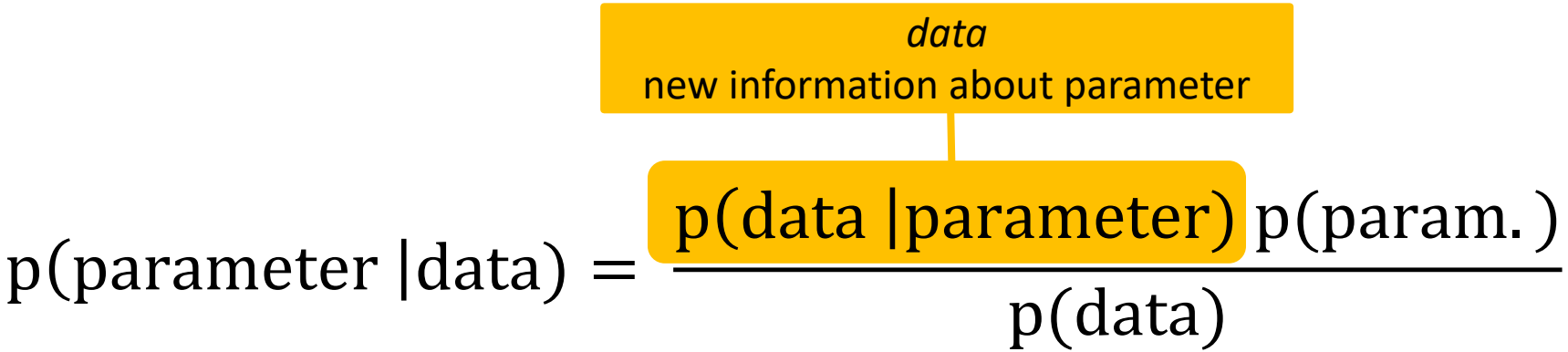
$$p(\text{parameter} \mid \text{data}) = \frac{p(\text{data} \mid \text{parameter})\, p(\text{param.})}{p(\text{data})}$$

BAYESIAN STATS

prior distribution
best guess before we see data

$$p(\text{parameter}\,|\text{data}) = \frac{p(\text{data}\,|\text{parameter})\,p(\text{param.})}{p(\text{data})}$$

BAYESIAN STATS



BAYESIAN & FREQUENTIST: Practically

- Both are valid statistical methods
- Both make *unverifiable* assumptions
  - Therefore, need to justify *specific* choice of methods
  - Are your assumptions reasonable, given scientific/stat knowledge?
- In simple cases, can define "equivalent" analyses
  - Simple linear model fit with either method will give similar slope est.
- In complex cases, with adequate data, usually consistent results
  - If not, you have a bigger problem than Bayesian or Frequentist: probably not enough data


FDA
U.S. FOOD & DRUG
ADMINISTRATION