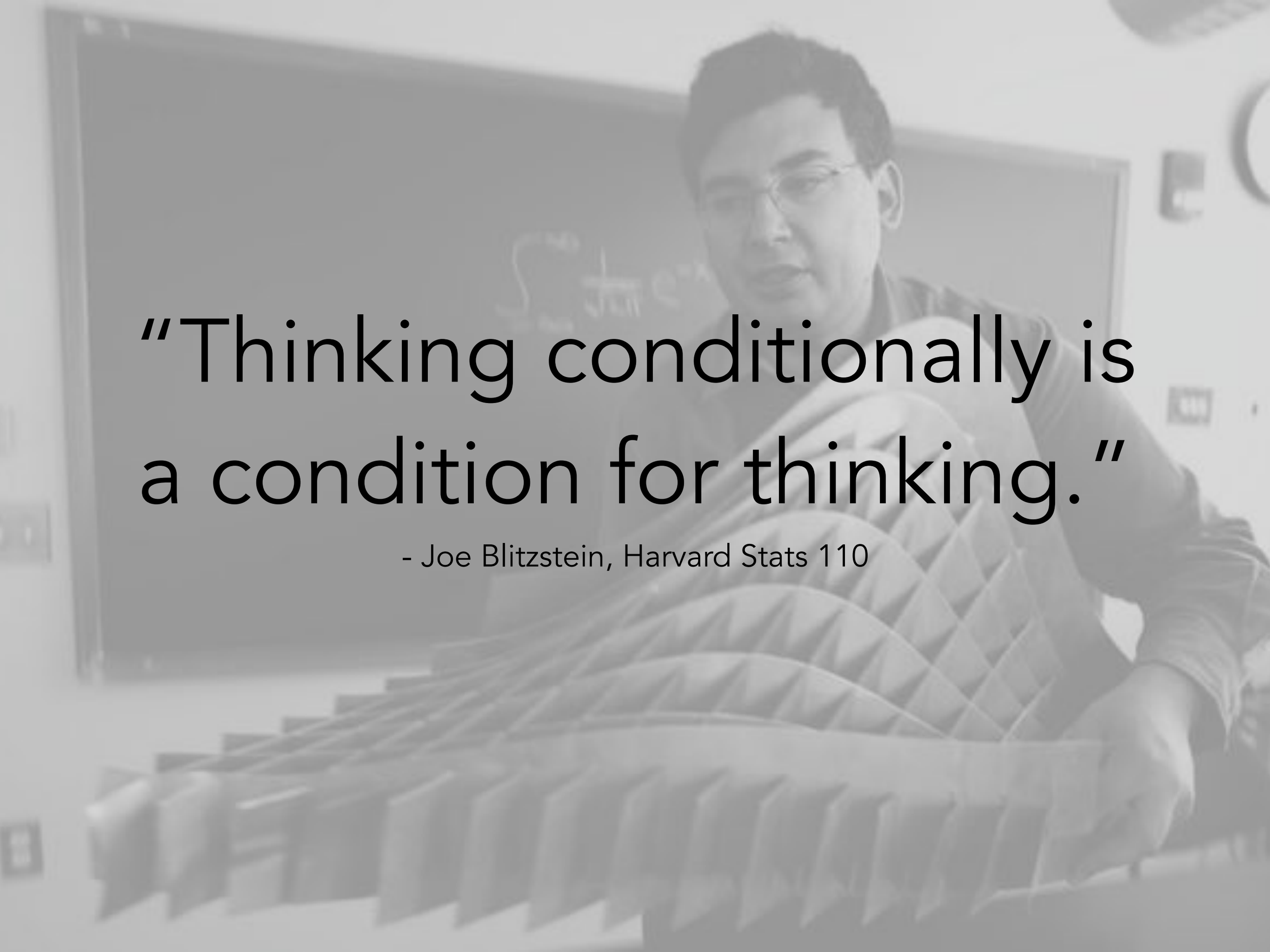
Thinking Conditionally

Matt Dickenson

Jan. 13, 2016



"Thinking conditionally is a condition for thinking."

- Joe Blitzstein, Harvard Stats 110

Outline

- Bayesian vs. Classical Statistics
- History of Bayesian Statistics
- Examples
- Benefits of Bayesian Methods
- Takeaway Lessons

- Classical ("frequentist") statistics
  - Many iterations
  - Algorithms applied to data
  - t-tests, p-values, statistical significance
- Bayesian statistics
  - Quantifying uncertainty
  - Probability distributions over beliefs
  - Priors, posteriors, decision functions

How tall is the Colorado State Capitol Building?

272′

Which step is exactly 5,280 feet above sea level?

1894: 15th

1969: 18th

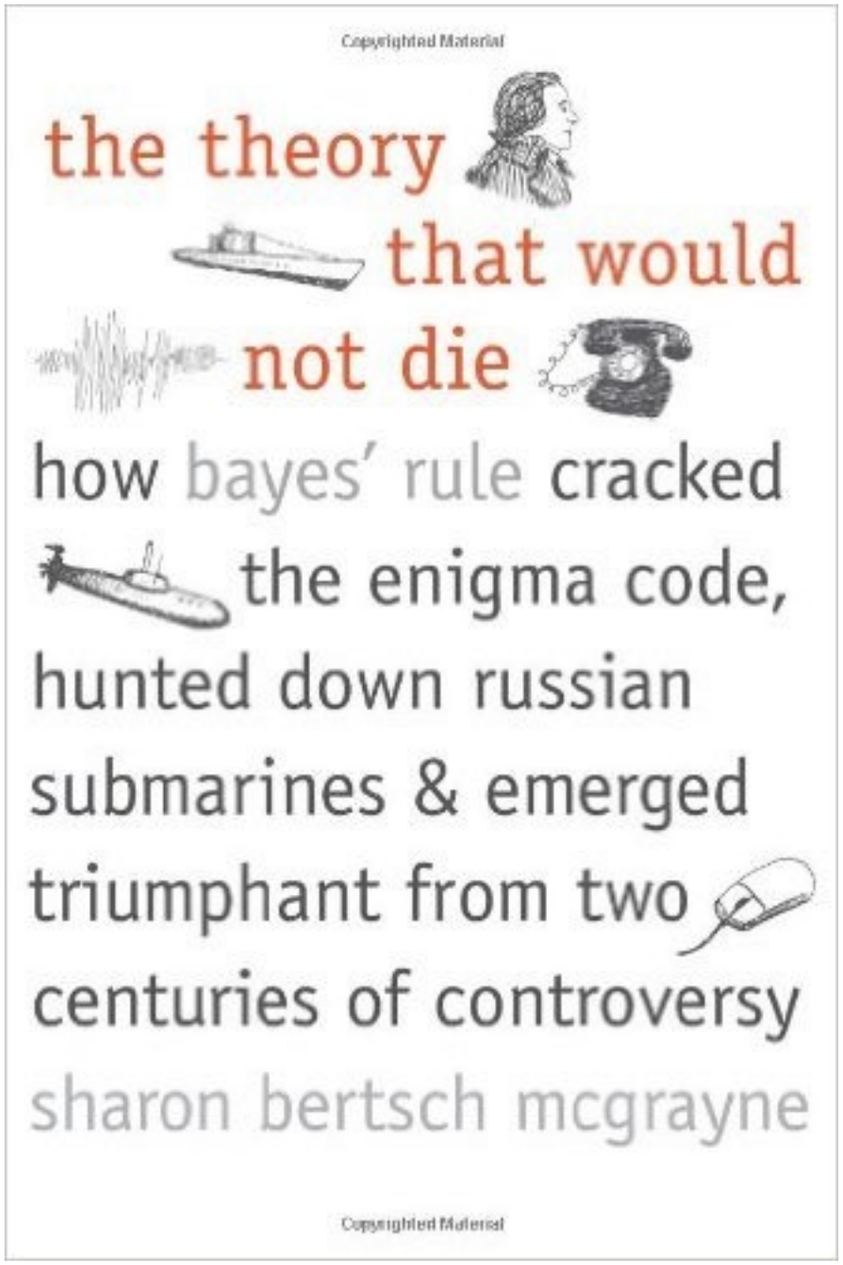
2003: 13th

2016: …?

History of Bayesian Thinking

- Thomas Bayes, 1701 - 1761 (Bayes Theorem)
- Pierre LaPlace, 1749-1827 (Generalized Bayes)
- R.A. Fisher, 1890 - 1962 (Randomization)
- Egon Pearson, 1895-1980 (Hypothesis Testing)
- Nicholas Metropolis, W.K. Hastings, Donald & Stuart Geman, Alan Gelfand, Donald Rubin, et al

History of Bayesian Thinking



Bayes' Rule

- P(A & B) = P(A|B) P(B) = P(B|A) P(A)
- P(A|B) = P(B|A) P(A) / P(B)

Coin Flips

- You find a quarter on the sidewalk and start flipping it. In the first three flips, it comes up heads twice and tails once. What odds would you be willing to accept that it will come up tails on the fourth flip?

- Classical statistician: 1:1
- Likelihood statistician: 2:1
- Bayesian statistician: 3:2

Fire Alarm

- You live in a building where the fire alarm goes off about once per week. There has never been a fire. The fire alarm is now going off. Should you evacuate?

- Classical statistician: No, the historical rate of fire is zero.

- Likelihood statistician: Same answer.

- Bayesian statistician: There's probably not a fire but getting stuck in a burning building is way worse, so yes!

Medical Test Example

- What is the probability that a patient has an illness, given a positive test result?
  - 1% of the population has a rare disease: P(disease) = 0.01
  - 99% of people who are sick test positive
  - 99% of people who are not sick test negative

Medical Test Example

- What's the probability of having the disease given a positive test result?
- P(Disease | Pos) = P(Pos | Disease) P(Disease) / P(Pos)
- = (99/100) (100/10,000) / (198/10,000) = 99/198 = 0.5

	Disease	Healthy	Total
Test Pos	99	99	198
Test Neg	1	9801	9802
Total	100	9900	10000

A/B Testing

- Suppose we are running an A/B test on a new feature.
- Experiment for 20 days, 10,000 impressions/day
- Require a “statistically significant” result to make a change
- Feature has no true effect—the clickthrough rate is always .1%.

A/B Testing

- If we stop the test every time we get a “significant” result ($p < .05$): 23% of tests suggest changing

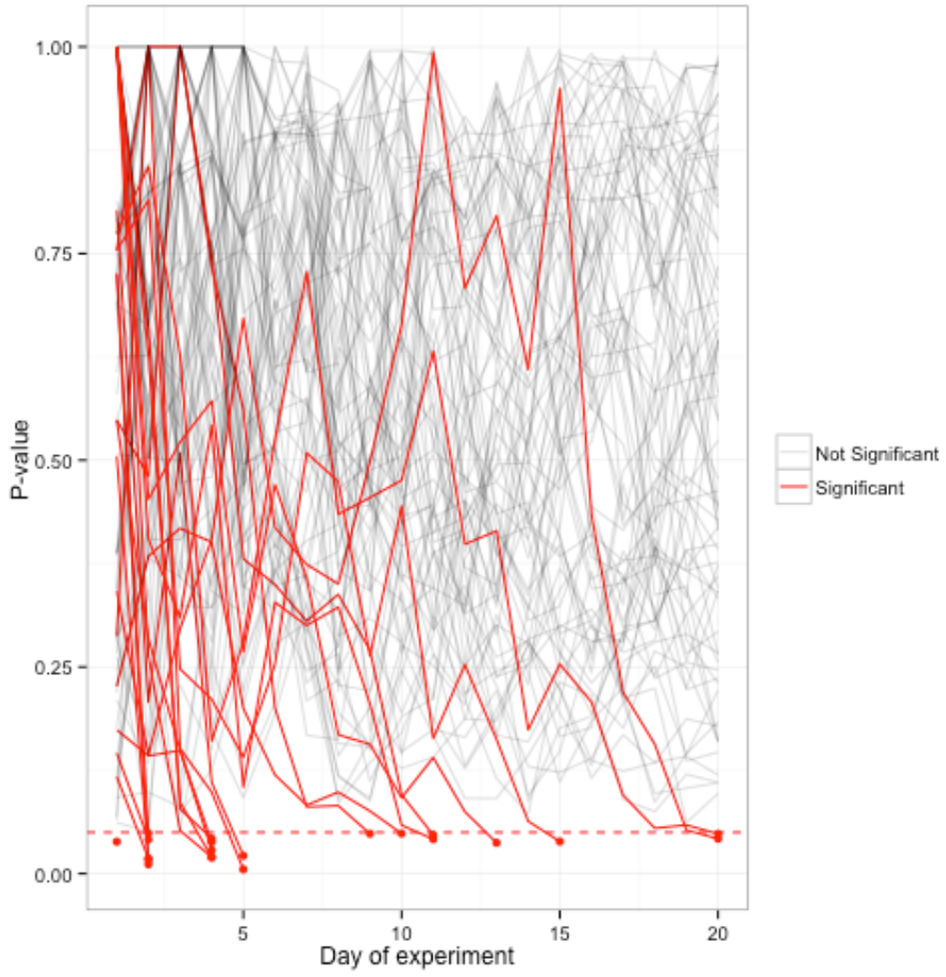


Image: varianceexplained.org

A/B Testing

- If we stick to the 20-day window, 5% of tests suggest changing

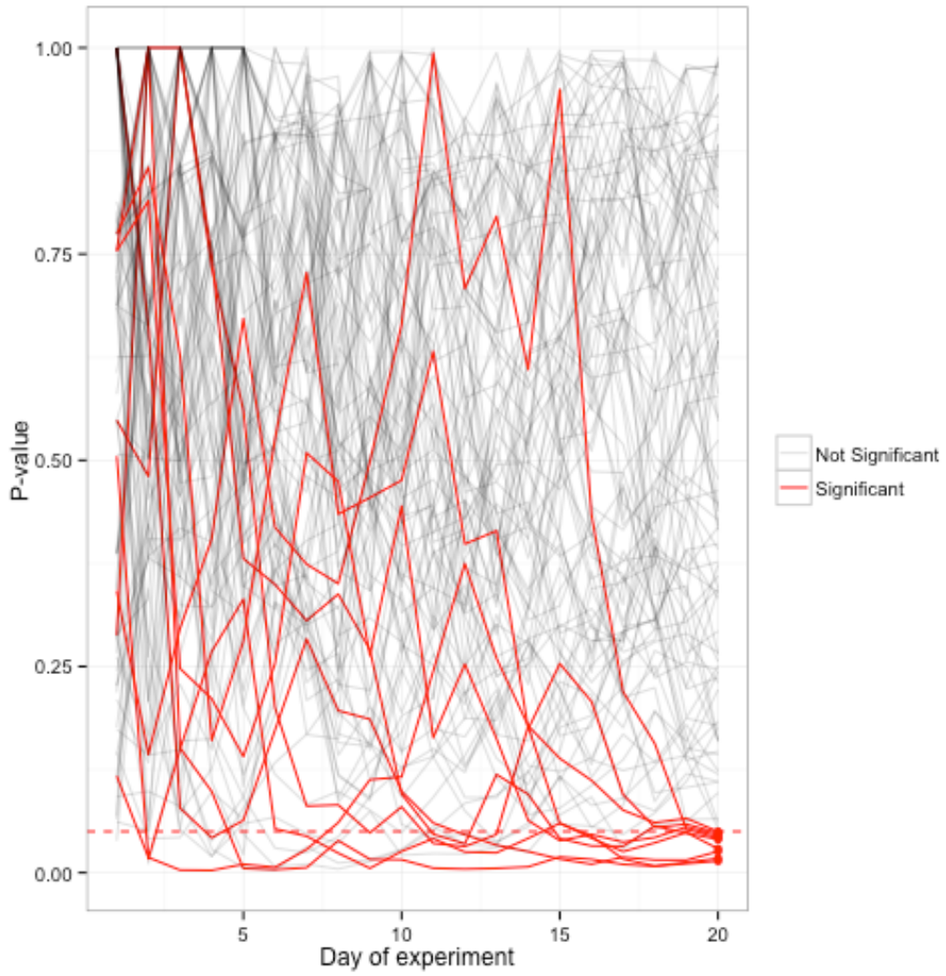


Image: varianceexplained.org

A/B Testing

- Bayesian alternatives:
  - At again given point in time, we have a probability distribution over the click-through rates in A & B
  - Estimate “expected loss” in both cases
  - Use a decision function
  - All prior data about “A” is also eligible for use

Benefits

- Sufficient statistics
- Online learning
- Avoids the “peeking problem”

Takeaways

- Conditional means
- Selecting on the dependent variable
- Survivorship bias

Resources

- **Books**
    - *The Theory that Wouldn't Die* by Sharon McGrayne
    - *The Signal and The Noise* by Nate Silver
    - *Think Bayes* by Allen Downey
    - *Bayesian Data Analysis* by Andrew Gelman, John Carlin, Hal Stern, David Dunson, Aki Vehtari, and Donald Rubin
- **Podcasts**
    - "Not so Standard Deviations"
    - "Planet Money" (ep 669)
- **Websites**
    - *nerds.airbnb.com*
    - *varianceexplained.org*