Bayesian Data Analysis
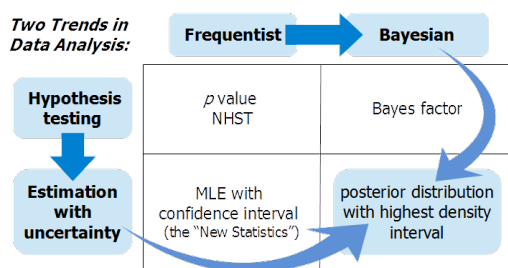## Getting Started

Min-Te Sun, Ph.D.

1

Textbooks

- Primary
  - Osvaldo Martin, "Bayesian Analysis with Python: Introduction to statistical modeling and probabilistic programming using PyMC3 and ArviZ," 2nd Ed., 2018.
  - Packt Publishing.
  - ISBN-13: 978-1789341652
- Reference 1
  - A. Gelman, J. B. Carlin, and H. S. Stern, "Bayesian Data Analysis," 3rd Ed., 2013.
  - Chapman & Hall/CRC Texts in Statistical Science
  - ISBN-13: 978-1439840955
- Reference 2
  - J. K. Kruschke, "Doing Bayesian Data Analysis - A Tutorial with R, JAGS, and Stan," 2nd Ed., 2015.
  - Academic Press.
  - ISBN-13: 978-0-12-405888-0

2

Why Bayesian Data Analysis?

Two Trends in Data Analysis:

	Frequentist	Bayesian
Hypothesis testing	p value NHST	Bayes factor
Estimation with uncertainty	MLE with confidence interval (the "New Statistics")	posterior distribution with highest density interval

Copyright © 2015 John K. Kruschke

3

Our Focus

- How to actually *do* Bayesian data analysis.
- Math concepts will be discussed (so you can gain the intuition of them) but we will not spend too much time on proving theorems.
- For those who are interested in rigorous proofs, please refer to the Reference 1 whenever necessary.

4

Prerequisite

- Really not much…
- Calculus => $\int x\,dx = \frac{1}{2}x^2$
- Basic programming experience is a big plus.
  - You need to learn the basics of Python, which is one of the most popular scripting languages.
  - You will need Python 3.6, PyMC3, and ArviZ.

5

Software Installation

- Python 3.6
  - Anaconda is recommended (https://www.anaconda.com/download/)
- PyMC3
  - "conda install -c conda-forge pymc3"
- ArviZ
  - "pip install arviz"
- An alternative way to install the necessary packages, once Anaconda is installed, is to go to https://github.com/aloctavodia/BAP and download the environment file named bap.yml. Using it, you can install all the necessary packages by doing the following:
 - "conda env create -f bap.yml"

6

What is Bayesian Inference?

- Reallocation of Credibility across Possibilities
- Credibility = Probability
- It is really about common sense!
  - When you see the floor is wet, what can you infer? (A) rain (B) drink spilled (C) dog pees
- "When you have eliminated the impossible, whatever remains, however improbable, must be the truth!" - Sherlock Holmes (or 柯南)
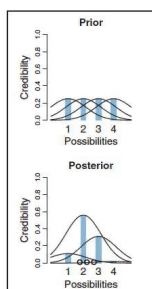
7

Sherlock Holmes (or 柯南) in Action
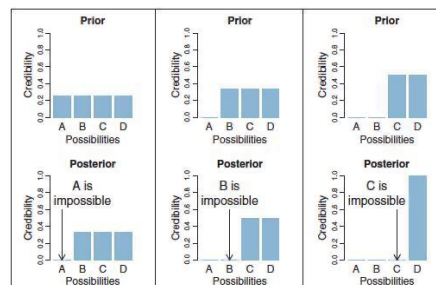
8

In Reality…

- Data are noisy
- Inferences are probabilistic

9

Possibilities : Parameter Values in Descriptive Models

- A key step in Bayesian analysis is defining the set of possibilities over which credibility is allocated.
  - Non-trivial
- Consider the example of the blood-pressure drug, in which blood pressures are measured in one group that took the drug and in another group that took a placebo
  - How much differences between these two groups will be considered the drug is effective?
  - *Our goal is to assess which possible descriptions are more or less credible*

10