An introduction to probabilistic programming with PyMC3

Sean Meling Murray and Solveig Masvie
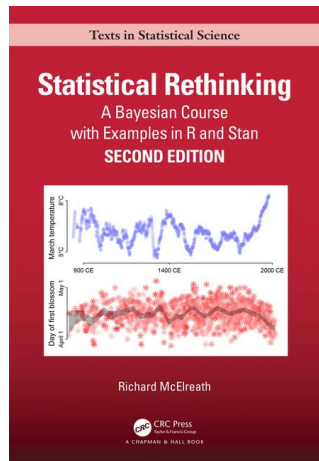
8. februar 2020

Table of contents

inmeta

Road map

- Theory
  - The basics of Bayesianism
  - Markov chain Monte Carlo methods (MCMC)
- Practice
  - Probabilistic programming with PyMC3

inmeta

What is Bayesian data analysis?

"A Bayesian is one who, vaguely expecting a horse, and catching a glimpse of a donkey, strongly believes he has seen a mule."

inmeta

What is Bayesian data analysis?

- Richard McElreath: "Bayesian inference is just counting."
- Count all the ways observed data could have arisen according to assumptions
- Assumptions that can arise in more ways are more consistent with the data, and therefore more plausible

Texts in Statistical Science

Statistical Rethinking
A Bayesian Course with Examples in R and Stan
SECOND EDITION

Richard McElreath

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

inmeta

The Frequentist vs. Bayesian debacle

- Frequentist statistics
  - ▶ Probability defined as the limiting frequency at which events occur
  - ▶ Uncertainty arises from sampling variation
- Bayesian statistics
  - ▶ Frequency and probability are different things
  - ▶ Uncertainty arises from our ignorance of the true state of the world

inmeta
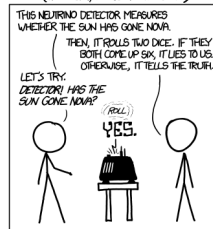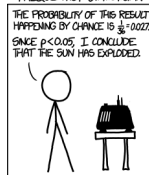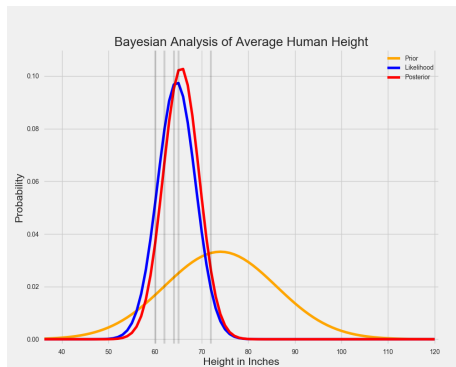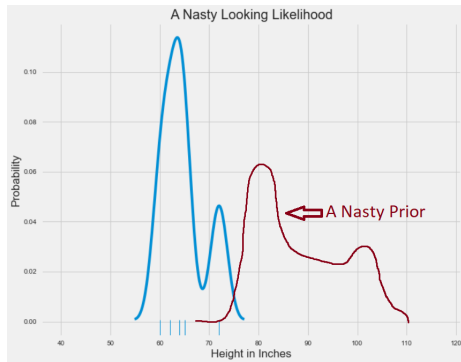
Bayesian Analysis

The prior distribution combined with likelihood distribution (observed data) equals posterior distribution

Bayesian Analysis of Average Human Height

inmeta

Bayesian Analysis

How do we find the posterior when the prior and likelihood distribution are complicated

A Nasty Looking Likelihood

A Nasty Prior

inmeta

Monte Carlo Simulations

Simple definition

- ▶ Monte Carlo simulations are just a way of estimating a fixed parameter by repeatedly generating random numbers.

The basis of Monte Carlo simulations is the Law of Large Numbers:

- ▶ As the number of identically distributed, randomly generated variables increases, their sample mean (average) approaches their theoretical mean.
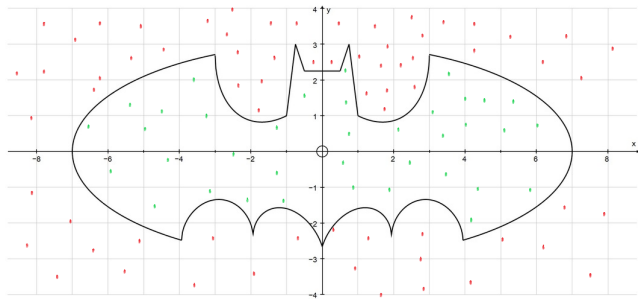
inmeta

Monte Carlo Simulations - Example

Use case: Calculate the area of this bat sign:

inmeta
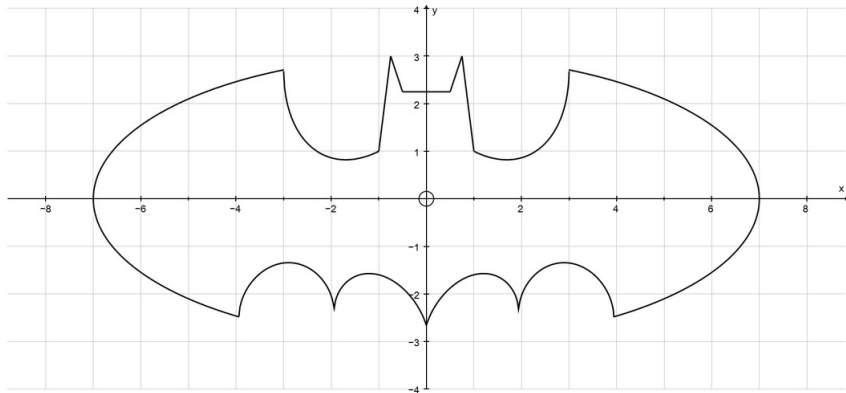
Monte Carlo Simulations - Example

$$\text{area of bat sign} = \text{area of rectangle} \times \frac{\text{green dots}}{\text{all dots}} \qquad (1)$$

inmeta

Markov Chain

- Sequences of events that have a probabilistic relation to one another
- Markov chains are memoryless. All we need to calculate the next event are available in the current state

inmeta

A slide with a theorem and a proof.

Theorem (Integral)

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

Bevis.

Here's the proof. □

inmeta

A slide with blocks

title of the bloc

bloc text

title of the bloc

bloc text

inmeta

A slide using pause

- Represent Abelian groups on the computer

inmeta

A slide using pause

- Represent Abelian groups on the computer
- Compute on Abelian groups

inmeta

A slide using pause

- Represent Abelian groups on the computer
- Compute on Abelian groups
- Solve equations, factor group homomorphisms

inmeta