An introduction to probabilistic programming with PyMC3

Sean Meling Murray and Solveig Masvie

10. februar 2020

Table of contents

Introduction

Bayesian statistics

Markov Chains

Road map

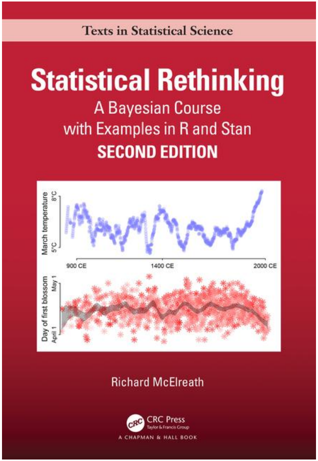
- Theory
    - The basics of Bayesianism
    - Markov chain Monte Carlo methods (MCMC)
- Practice
    - Probabilistic programming with PyMC3

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."

What is Bayesian data analysis?

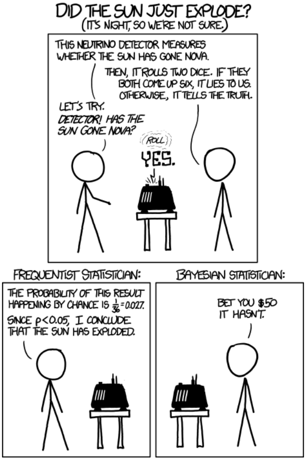
- Richard McElreath: “Bayesian inference is just counting.”
- Count all the ways observed data could have arisen according to assumptions
- Assumptions that can arise in more ways are more consistent with the data, and therefore more plausible



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

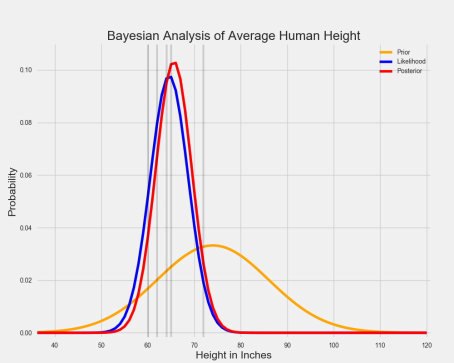


Bayesian Analysis

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

$$P(\theta|D,M) = \frac{P(D|\theta,M)P(\theta|M)}{P(D|M)} \tag{1}$$

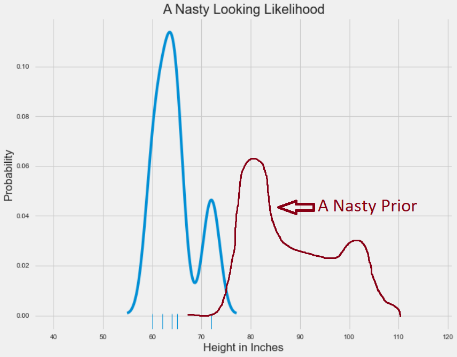
$$\mathscr{P} \propto \mathscr{L}\Pi \tag{2}$$



inmeta

Bayesian Analysis

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



inmeta

Monte Carlo Simulations

Simple definition

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

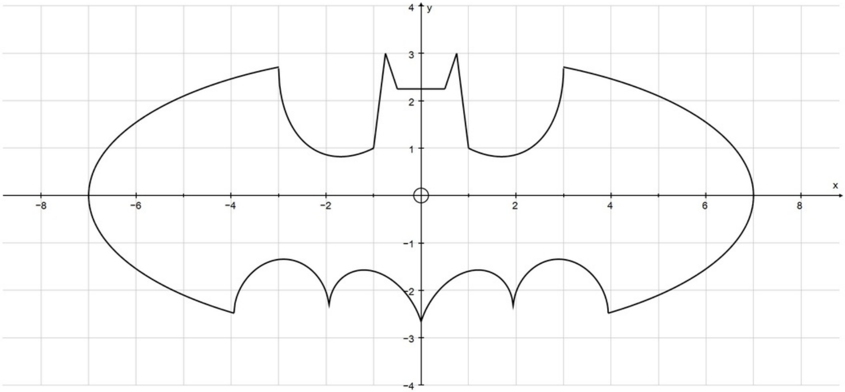
The basis of Monte Carlo simulations is the Law of Large Numbers:

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

$$\frac{1}{N}\sum_{i=1}^{N} Z_i \to E[Z],\ N \to \infty \tag{3}$$

Monte Carlo Simulations - Example

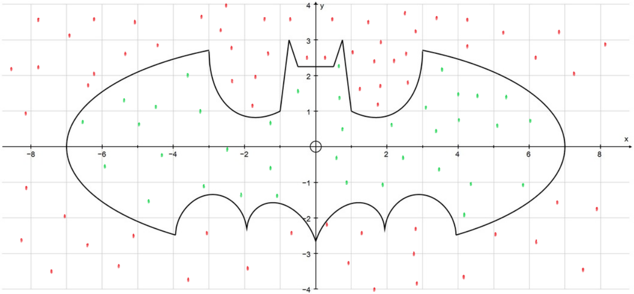
Use case: Calculate the area of this bat sign:



inmeta

Monte Carlo Simulations - Example

$$\text{area of bat sign} = \text{area of rectangle} \times \frac{\text{green dots}}{\text{all dots}} \tag{4}$$



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

Bayesian MCMC

Steps in MCMC

- Define function for $\mathscr{L}$, Π and thus $\mathscr{P}$
- Define initial guess for θ (based on the prior)
- Try a jump in θ
- Accept/reject based on chosen method/sampler (Metropolis)
- Keep jumping
- After doing many steps remove burn-in steps

Priors

It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so. Mark Twain

Likelihood

Posterior

Sampler

Metropolis rule

- If $\mathscr{P}_{new} > \mathscr{P}_i$ accept the jump so $\theta_{i+1} = \theta_i$
- If $\mathscr{P}_{new} < \mathscr{P}_i$ accept the jump with probability $\frac{\mathscr{P}_{new}}{\mathscr{P}_i}$

inmeta

Jump in θ

How to make a jump in θ

$$\theta_{new} = \theta_i + \mathcal{N}(0, \Delta\theta) \tag{5}$$

We call $\Delta\theta$ for the jump scale. Normally this must be tuned manually for every dimension. A rule of thumb is that we want a jump scale that gives a reasonable acceptance rate.

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. $\square$

inmeta

A slide with blocks

title of the bloc

bloc text

title of the bloc

bloc text

A slide using pause

- Represent Abelian groups on the computer

A slide using pause

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

A slide using pause

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