
Bayesian shenaniganry
by Justin and Stephen

Goals

Bayesian theory

Probabilistic reasoning and Bayes' theorem

Bayesian parameter inference (in 1- and 2-d)

Choosing priors

Bayesian computation

aka "fitting the fudge out of a straight line"

MCMC sampling (MH, EMCEE, Gibbs, Stan)

Graphical (hierarchical) models

Probabilistic reasoning

Theory Empirical

deductive reasoning

theory / hypotheses → *predictions* → *observations (uncertainty/ incomplete knowledge)*

inductive or "plausible" reasoning

Probability

~~Plausibility~~ of hypothesis $\mathcal{H}$ given some observations d any prior information $\mathcal{I}$

$$P(\mathcal{H}|\mathrm{d},\mathcal{I})$$

Cox (1946): What are the quantitative rules for logical and consistent inductive reasoning? ***Answer: Probability theory***

Probabilistic reasoning (cheat sheet)

(scale)

$$P(\text{true}) = 1,\ P(\text{false}) = 0$$

(sum rule)

$$P(X|I) + P(\bar{X}|I) = 1$$

(product rule)

$$P(X, Y|I) = P(X|Y, I)P(Y|I) = P(Y|X, I)P(X|I)$$

$$\Rightarrow \boxed{P(\mathcal{H}|\mathbf{d}, I) = \frac{P(\mathbf{d}|\mathcal{H}, I)P(\mathcal{H}|I)}{P(\mathbf{d}|I)}}$$

(Bayes' theorem)

(normalization)

$$\int_{\text{all } X} P(X|I)dX = 1$$

(marginalization)

$$P(X|I) = \int_{\text{all } Y} P(X, Y|I)dY$$

(change of variables)

$$P(Y(X)) = P(X)\left|\frac{dX}{dY}\right|$$

(independence)

$$P(X|Y) = P(X)$$

$$\Rightarrow P(X, Y) = P(X)P(Y)$$

Probabilistic reasoning (exercises)

Dude, are you infected?

A nasty disease breaks out and infects 1% of the population, with no cure and certain nasty death. A test is developed that has 99% reliability; that is, 99% of people who are sick test positive and 99% of the healthy people test negative. You go to get tested and get a positive; how worried should you be?

Monty Hall Problem

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to switch to door No. 2?" Is it to your advantage to switch?

Bayesian parameter inference

Model, $\mathcal{M}$ ***Parameters,*** θ ***Data,*** d_o ***Prior information,*** $\mathcal{I}$

posterior

$$\boxed{P(\boldsymbol{\theta}|\mathbf{d}_o, \mathcal{M}, \mathcal{I})}$$

"Scientific process"

1. ***Theory + Experiment design = Model*** $\mathcal{M}$ ***with parameters*** θ
+ Prior information $\mathcal{I}$ $\Rightarrow P(\boldsymbol{\theta}|\mathcal{M}, \mathcal{I})$ ***prior***

2. ***Make predictions (aka, "deductive reasoning") from model***

$$\Rightarrow P(\mathbf{d}|\boldsymbol{\theta}, \mathcal{M}, \mathcal{I})$$ ***sampling distribution***

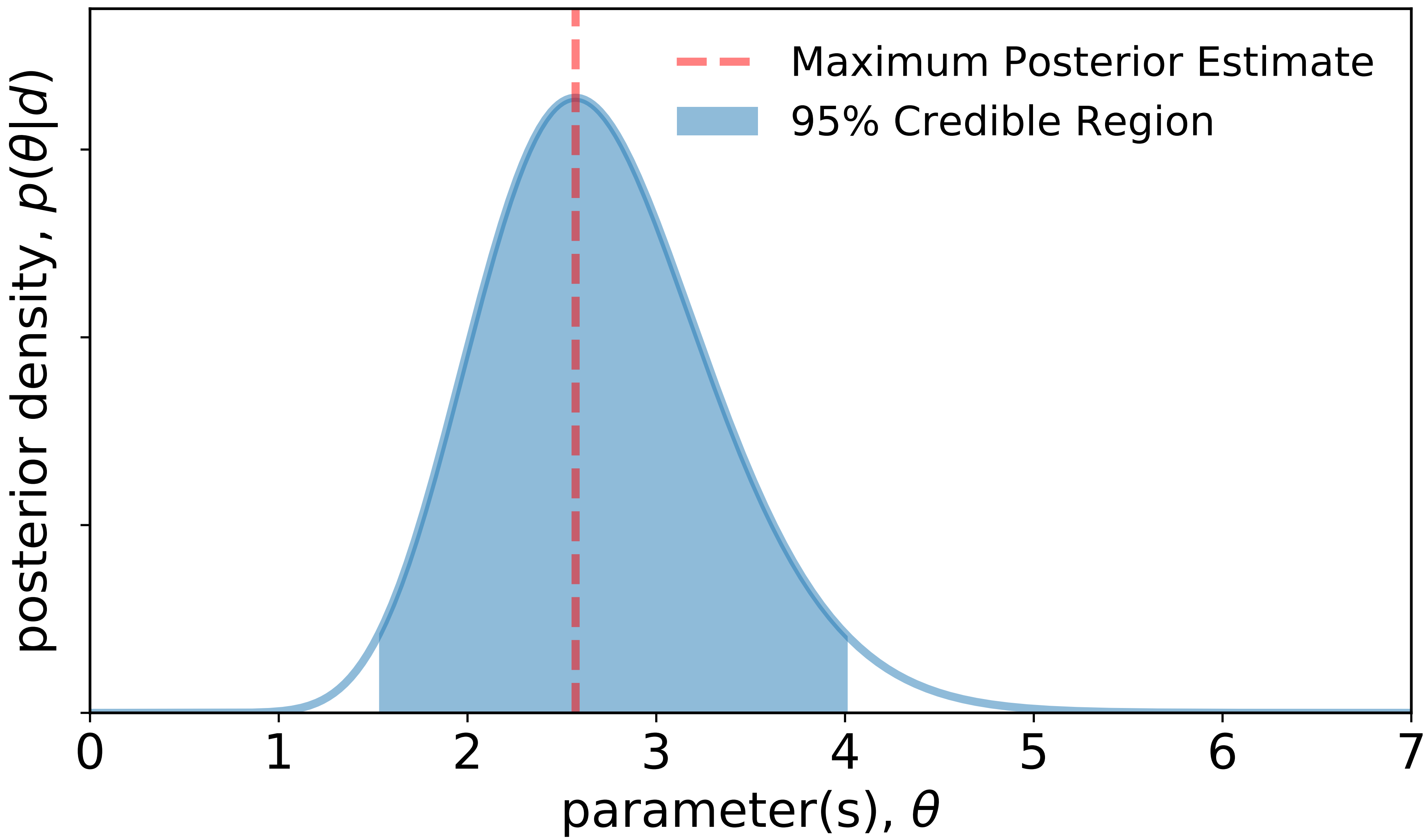
3. ***Do experiment*** $\mathrm{d}_o \Rightarrow P(\mathbf{d}_o|\boldsymbol{\theta}, \mathcal{M}, \mathcal{I})$ ***likelihood function***

4. ***Inference (aka, "inductive reasoning")*** ***posterior***

$$\boxed{P(\boldsymbol{\theta}|\mathbf{d}_o, \mathcal{M}, \mathcal{I})} = \frac{P(\mathbf{d}_o|\boldsymbol{\theta}, \mathcal{M}, \mathcal{I}) \times P(\boldsymbol{\theta}|\mathcal{M}, \mathcal{I})}{P(\mathbf{d}_o|\mathcal{M}, \mathcal{I})}$$

Bayesian parameter inference

Model, $\mathcal{M}$ Parameters, θ Data, d_o Prior information, $\mathcal{I}$

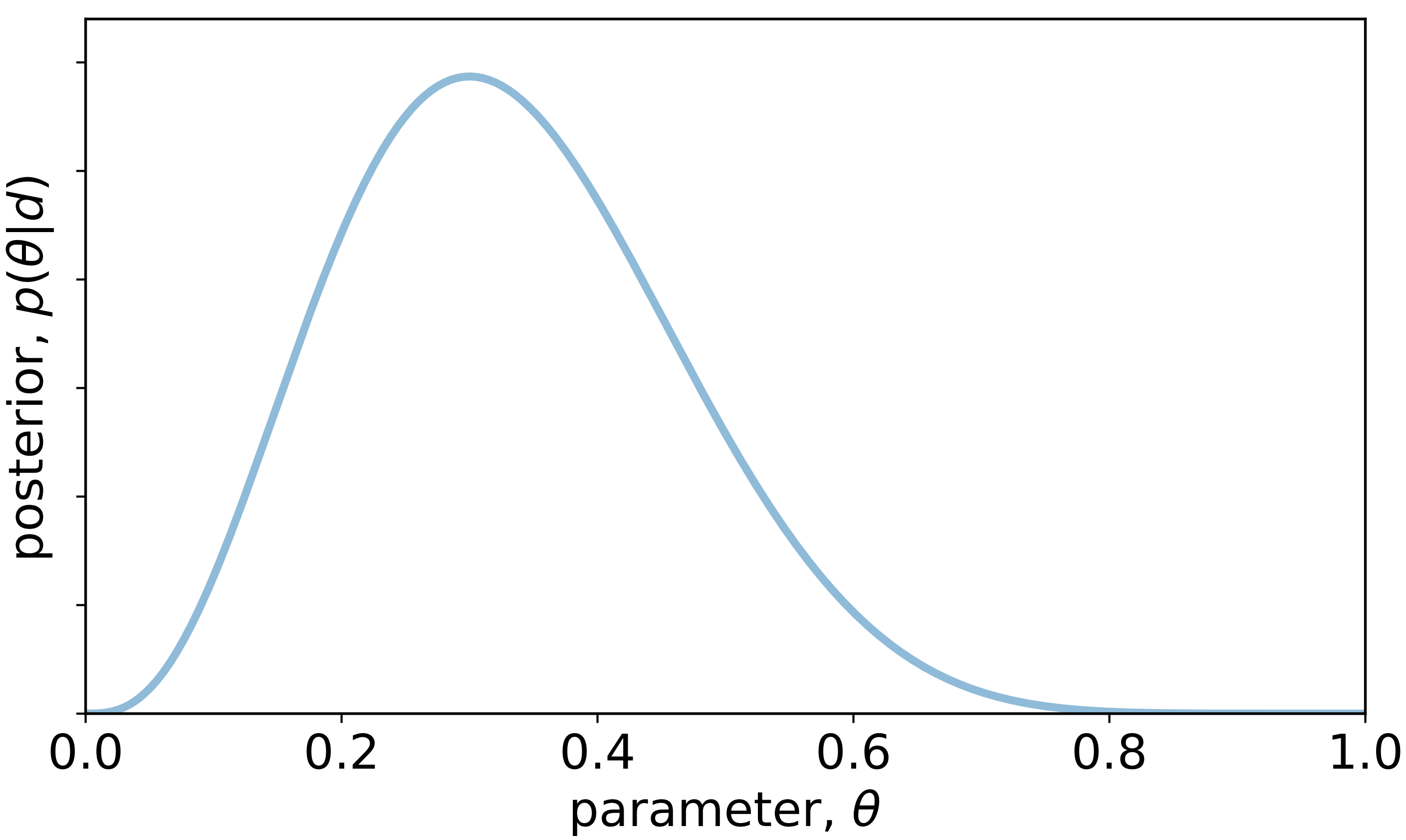


$$P(\boldsymbol{\theta}|\mathrm{d}_o, \mathcal{M}, \mathcal{I}) = \frac{P(\mathrm{d}_o|\boldsymbol{\theta}, \mathcal{M}, \mathcal{I}) \times P(\boldsymbol{\theta}|\mathcal{M}, \mathcal{I})}{P(\mathrm{d}_o|\mathcal{M}, \mathcal{I})}$$

Bayesian parameter inference (exercises)

Coin toss

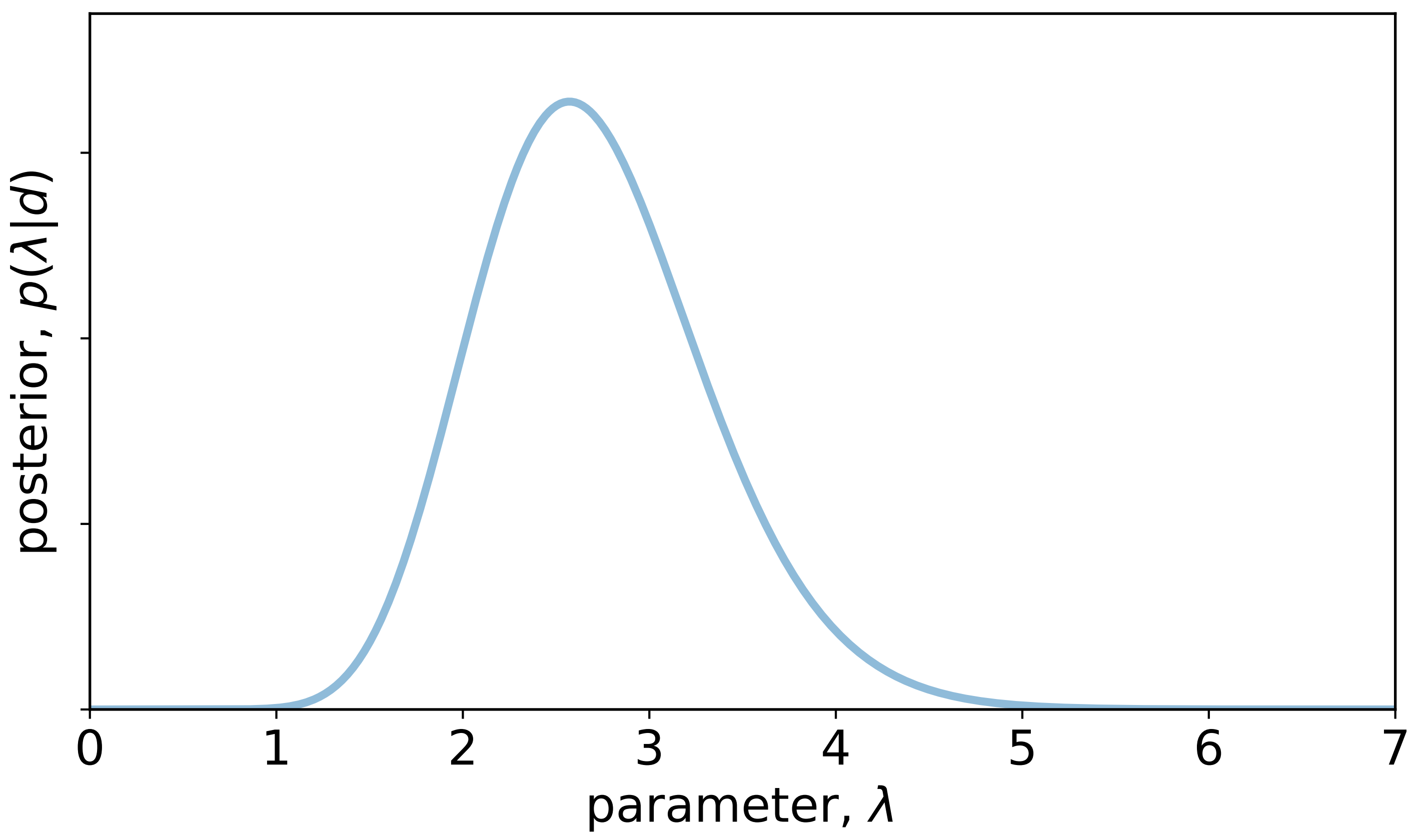
You have a coin and want to know the probability with which it gives heads, θ. You make 10 tosses and get 3H, 7T. What is the posterior for θ? (assume tosses are independent and uniform prior, for now)



Bayesian parameter inference (exercises)

Decay rate

You have some nuclear material and want to know the decay rate, λ (per day). You measure the number of decays for seven days and get d = (4, 0, 3, 2, 4, 2, 3). What's your posterior for λ? (assume decays are independent and Poisson, and uniform prior for now)



Bayesian parameter inference (exercises)

How heavy your baby?

You're at the doctor getting your baby weighed. The doc takes five measurements: d = (4.49, 5.15, 5.26 , 5.90, 4.86) kg. On the scale it (conveniently) says it has uncorrelated Gaussian errors with σ=0.5kg. What's your posterior for your baby's mass? (assume uniform priors)

Hold your hockey sticks, I don't trust that label

You're skeptical that the stated uncertainties on the scales are correct. What's your joint posterior for your baby's mass and the std deviation of the scales σ?

Bonus

Can you marginalize over σ and compare with the case of known std deviation? (Spoiler: you should get a Student's-t posterior for your baby's mass)

Choosing priors (cheat sheet)

There is no such thing as a truly "uninformative prior"

(even an infinite uniform prior contains information!)

"uniform = uninformative" is <u>not</u> a consistent rule for choosing prior

(when you change variables, it all goes to…)

Physics ***vs.*** **Information theory** ***vs.*** **Convenience**

Jeffrey's priors (in 1D)

$$P_J(\boldsymbol{\theta}) \propto \sqrt{\mathcal{F}(\boldsymbol{\theta})} \qquad \mathcal{F}(\boldsymbol{\theta}) = \left\langle -\frac{\partial^2 \ln P(\mathbf{d}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \right\rangle_{P(\mathbf{d}|\boldsymbol{\theta})}$$

Conjugate priors (aka, "keeping it in the family")

Gaussian

$$e^{-\frac{1}{2}(\theta-\mu)^2/\sigma^2}$$

Gamma

$$\theta^{\alpha-1}e^{-\beta\theta}$$

Inverse-Gamma

$$\theta^{-(\alpha+1)}e^{-\beta/\theta}$$

Beta

$$\theta^{\alpha-1}(1-\theta)^{\beta-1}$$

https://en.wikipedia.org/wiki/Conjugate_prior

Choosing priors (exercises)

What are the Jeffrey's priors for the coin toss, nuclear decay, and weighing the baby problems?

Bonus: plot them alongside the likelihood functions, show how they change the posteriors compared to uniform priors

What are the conjugate priors for those three problems? Convince yourself that the Jeffrey's priors are limiting cases of the conjugate priors