Conjugate Priors: Beta and Normal

18.05 Spring 2014

Review: Continuous priors, discrete data

'Bent' coin: unknown probability θ of heads.

Prior $f(\theta) = 2\theta$ on [0,1].

Data: heads on one toss.

Question: Find the posterior pdf to this data.

hypoth.	prior	likelihood	Bayes numerator	posterior
θ	$2\theta\, d\theta$	θ	$2\theta^2\, d\theta$	$3\theta^2\, d\theta$
Total	1		$T = \int_0^1 2\theta^2\, d\theta = 2/3$	1

Posterior pdf: $f(\theta|x) = 3\theta^2$.

Review: Continuous priors, continuous data

Bayesian update table

hypoth.	prior	likeli.	Bayes numerator	posterior
θ	$f(\theta)\,d\theta$	$f(x\,\vert\,\theta)$	$f(x\,\vert\,\theta)f(\theta)\,d\theta$	$f(\theta\,\vert\,x)\,d\theta = \dfrac{f(x\,\vert\,\theta)f(\theta)\,d\theta}{f(x)}$
total	1		$f(x)$	1

$$f(x) = \int f(x\,|\,\theta)f(\theta)\,d\theta$$

Notice that we overuse the letter f. It is a generic symbol meaning 'whatever function is appropriate here'.

Romeo and Juliet

See class 14 slides

Updating with normal prior and normal likelihood

A normal prior is conjugate to a normal likelihood with known σ.

- Data: $x_1, x_2, \ldots, x_n$
- **Normal likelihood.** $x_1, x_2, \ldots, x_n \sim \mathrm{N}(\theta, \sigma^2)$

  Assume θ is our unknown parameter of interest, σ is known.
- **Normal prior.** $\theta \sim \mathrm{N}(\mu_{\text{prior}}, \sigma^2_{\text{prior}})$.
- **Normal Posterior.** $\theta \sim \mathrm{N}(\mu_{\text{post}}, \sigma^2_{\text{post}})$.

- We have simple updating formulas that allow us to avoid complicated algebra or integrals (see next slide).

hypoth.	prior	likelihood	posterior
θ	$f(\theta) \sim \mathrm{N}(\mu_{\text{prior}}, \sigma^2_{\text{prior}})$	$f(x\|\theta) \sim \mathrm{N}(\theta, \sigma^2)$	$f(\theta\|x) \sim \mathrm{N}(\mu_{\text{post}}, \sigma^2_{\text{post}})$
θ	$c_1 \exp\left(\frac{-(\theta-\mu_{\text{prior}})^2}{2\sigma^2_{\text{prior}}}\right)$	$c_2 \exp\left(\frac{-(x-\theta)^2}{2\sigma^2}\right)$	$c_3 \exp\left(\frac{-(\theta-\mu_{\text{post}})^2}{2\sigma^2_{\text{post}}}\right)$

Board question: Normal-normal updating formulas

$$a = \frac{1}{\sigma^2_{\text{prior}}} \qquad b = \frac{n}{\sigma^2}, \qquad \mu_{\text{post}} = \frac{a\mu_{\text{prior}} + b\bar{x}}{a+b}, \qquad \sigma^2_{\text{post}} = \frac{1}{a+b}.$$

Suppose we have one data point $x = 2$ drawn from $\mathrm{N}(\theta, 3^2)$

Suppose θ is our parameter of interest with prior $\theta \sim \mathrm{N}(4, 2^2)$.

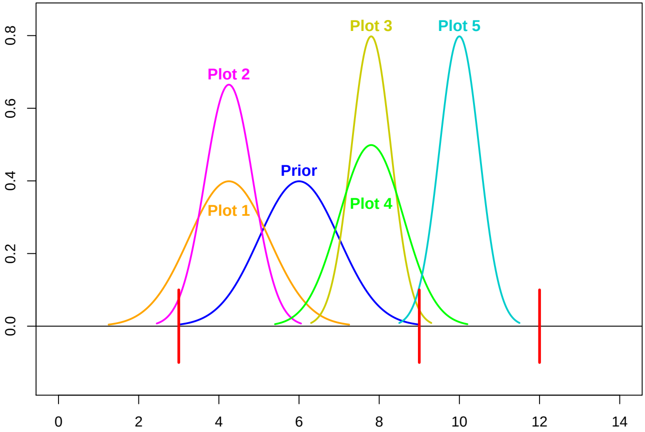
0. Identify μ_{prior}, σ_{prior}, σ, n, and $\bar{x}$.

1. Make a Bayesian update table, but leave the posterior as an unsimplified product.

2. Use the updating formulas to find the posterior.

3. By doing enough of the algebra, understand that the updating formulas come by using the updating table and doing a lot of algebra.

Concept question: normal priors, normal likelihood

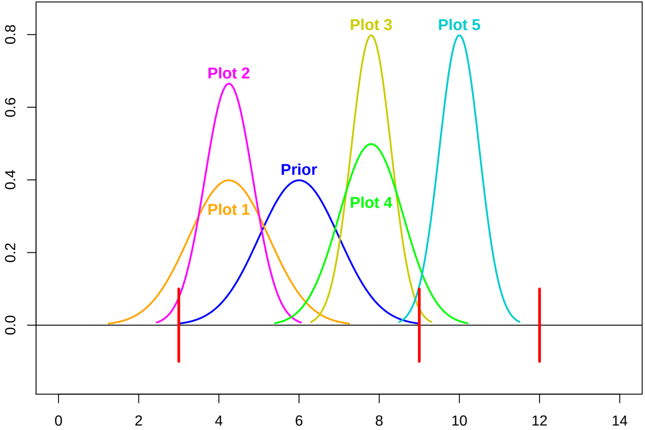


Blue graph = prior   Red lines = data in order: 3, 9, 12

(a) Which plot is the posterior to just the first data value?

(Click on the plot number.)

Concept question: normal priors, normal likelihood



Blue graph = prior Red lines = data in order: 3, 9, 12

(b) Which graph is posterior to all 3 data values?

(Click on the plot number.)

Board question: normal/normal

For data $x_1, \ldots, x_n$ with data mean $\bar{x} = \frac{x_1 + \ldots + x_n}{n}$

$$a = \frac{1}{\sigma^2_{\text{prior}}} \qquad b = \frac{n}{\sigma^2}, \qquad \mu_{\text{post}} = \frac{a\mu_{\text{prior}} + b\bar{x}}{a+b}, \qquad \sigma^2_{\text{post}} = \frac{1}{a+b}.$$

Question. On a basketball team the average free throw percentage over all players is a $\mathrm{N}(75, 36)$ distribution. In a given year individual players free throw percentage is $\mathrm{N}(\theta, 16)$ where θ is their career average.

This season Sophie Lie made 85 percent of her free throws. What is the posterior expected value of her career percentage θ?

Conjugate priors

A prior is conjugate to a likelihood if the posterior is the same type of distribution as the prior.
Updating becomes algebra instead of calculus.

	hypothesis	data	prior	likelihood	posterior
Bernoulli/Beta	$\theta \in [0,1]$	x	$\text{beta}(a,b)$	$\text{Bernoulli}(\theta)$	$\text{beta}(a+1,b)$ or $\text{beta}(a,b+1)$
	θ	$x=1$	$c_1\theta^{a-1}(1-\theta)^{b-1}$	θ	$c_3\theta^{a}(1-\theta)^{b-1}$
	θ	$x=0$	$c_1\theta^{a-1}(1-\theta)^{b-1}$	$1-\theta$	$c_3\theta^{a-1}(1-\theta)^{b}$
Binomial/Beta	$\theta \in [0,1]$	x	$\text{beta}(a,b)$	$\text{binomial}(N,\theta)$	$\text{beta}(a+x,b+N-x)$
(fixed N)	θ	x	$c_1\theta^{a-1}(1-\theta)^{b-1}$	$c_2\theta^{x}(1-\theta)^{N-x}$	$c_3\theta^{a+x-1}(1-\theta)^{b+N-x-1}$
Geometric/Beta	$\theta \in [0,1]$	x	$\text{beta}(a,b)$	$\text{geometric}(\theta)$	$\text{beta}(a+x,b+1)$
	θ	x	$c_1\theta^{a-1}(1-\theta)^{b-1}$	$\theta^{x}(1-\theta)$	$c_3\theta^{a+x-1}(1-\theta)^{b}$
Normal/Normal	$\theta \in (-\infty,\infty)$	x	$\text{N}(\mu_{\text{prior}},\sigma^2_{\text{prior}})$	$\text{N}(\theta,\sigma^2)$	$\text{N}(\mu_{\text{post}},\sigma^2_{\text{post}})$
(fixed σ^2)	θ	x	$c_1\exp\left(\frac{-(\theta-\mu_{\text{prior}})^2}{2\sigma^2_{\text{prior}}}\right)$	$c_2\exp\left(\frac{-(x-\theta)^2}{2\sigma^2}\right)$	$c_3\exp\left(\frac{(\theta-\mu_{\text{post}})^2}{2\sigma^2_{\text{post}}}\right)$

There are many other likelihood/conjugate prior pairs.

Concept question: conjugate priors

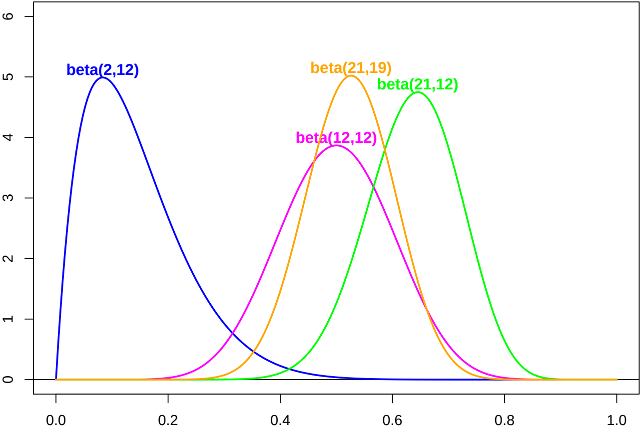
Which are conjugate priors?

	hypothesis	data	prior	likelihood
a) Exponential/Normal	$\theta \in [0, \infty)$	x	$\mathrm{N}(\mu_{\text{prior}}, \sigma^2_{\text{prior}})$	$\exp(\theta)$
	θ	x	$c_1 \exp\left(-\frac{(\theta-\mu_{\text{prior}})^2}{2\sigma^2_{\text{prior}}}\right)$	$\theta \mathrm{e}^{-\theta x}$
b) Exponential/Gamma	$\theta \in [0, \infty)$	x	$\mathrm{Gamma}(a, b)$	$\exp(\theta)$
	θ	x	$c_1 \theta^{a-1} \mathrm{e}^{-b\theta}$	$\theta \mathrm{e}^{-\theta x}$
c) Binomial/Normal	$\theta \in [0, 1]$	x	$\mathrm{N}(\mu_{\text{prior}}, \sigma^2_{\text{prior}})$	$\mathrm{binomial}(N, \theta)$
(fixed N)	θ	x	$c_1 \exp\left(-\frac{(\theta-\mu_{\text{prior}})^2}{2\sigma^2_{\text{prior}}}\right)$	$c_2\, \theta^x (1-\theta)^{N-x}$

1. none	2. a	3. b	4. c
5. a,b	6. a,c	7. b,c	8. a,b,c

Variance can increase

Normal-normal: variance **always** decreases with data.
Beta-binomial: variance **usually** decreases with data.



Variance of beta(2,12) (blue) is bigger than that of beta(12,12) (magenta), but beta(12,12) can be a posterior to beta(2,12)

Table discussion: likelihood principle

Suppose the prior has been set. Let x_1 and x_2 be two sets of data. Which of the following are true.

(a) If the likelihoods $f(x_1|\theta)$ and $f(x_2|\theta)$ are the same then they result in the same posterior.

(b) If x_1 and x_2 result in the same posterior then their likelihood functions are the same.

(c) If the likelihoods $f(x_1|\theta)$ and $f(x_2|\theta)$ are proportional then they result in the same posterior.

(d) If two likelihood functions are proportional then they are equal.

Concept question: strong priors

Say we have a bent coin with unknown probability of heads θ.

We are convinced that $\theta \leq 0.7$.
Our prior is uniform on $[0, 0.7]$ and 0 from 0.7 to 1.

We flip the coin 65 times and get 60 heads.

Which of the graphs below is the posterior pdf for θ?