Conjugate Priors: Beta and Normal; Choosing Priors

18.05 Spring 2014

Jeremy Orloff and Jonathan Bloom

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	unnormalized posterior	posterior
$\theta \pm \frac{d\theta}{2}$	$2\theta\, d\theta$	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 tables with and without infinitesimals

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

hypoth.	prior	likeli.	unnormalized posterior	posterior
$\theta \pm \frac{d\theta}{2}$	$f(\theta)\,d\theta$	$f(x\,\vert\,\theta)\,dx$	$f(x\,\vert\,\theta)f(\theta)\,d\theta\,dx$	$f(\theta\,\vert\,x)\,d\theta = \dfrac{f(x\,\vert\,\theta)f(\theta)\,d\theta\,dx}{f(x)\,dx}$
total	1		$f(x)\,dx$	1

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

Board question: Romeo and Juliet

Romeo is always late. How late follows a uniform distribution uniform$(0, \theta)$ with unknown parameter θ in hours.

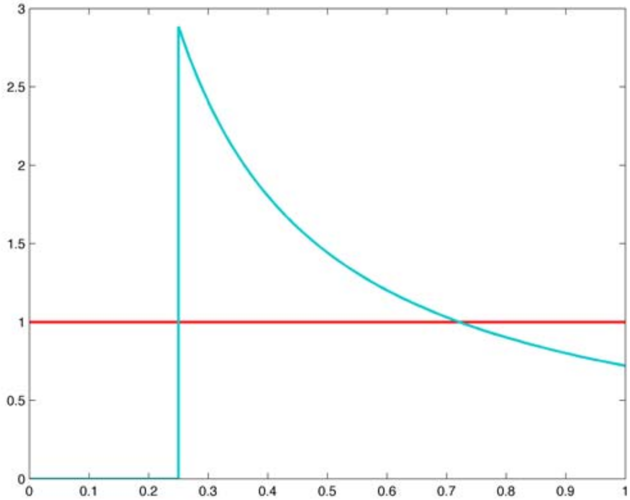
Juliet knows that $\theta \leq 1$ hour and she assumes a flat prior for θ on $[0, 1]$.

On their first date Romeo is 15 minutes late.

(a) find and graph the prior and posterior pdf's for θ

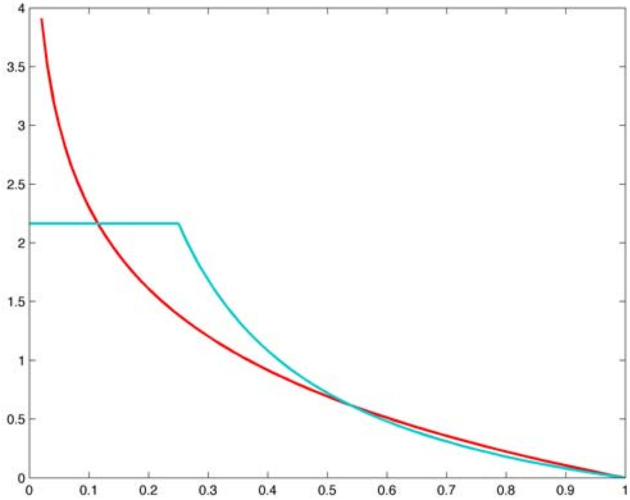
(b) find and graph the prior predictive and posterior predictive pdf's of how late Romeo will be on the second data (if he gets one!).

Solution: prior and posterior graphs



Prior and posterior pdf's for θ.

Solution: predictive prior and posterior graphs



Prior (red) and posterior (blue) predictive pdf's for x_2

Updating with normal prior and normal likelihood

- Data: $x_1, x_2, \ldots, x_n$ drawn from $\mathrm{N}(\theta, \sigma^2)/$
- Assume θ is our unknown parameter of interest, σ is known.
- Prior: $\theta \sim \mathrm{N}(\mu_{\text{prior}}, \sigma^2_{\text{prior}})$

In this case the posterior for θ is $\mathrm{N}(\mu_{\text{post}}, \sigma^2_{\text{post}})$ with

$$a = \frac{1}{\sigma^2_{\text{prior}}} \qquad b = \frac{n}{\sigma^2}, \qquad \bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

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

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. Use the updating formulas to find the posterior.

2. Find the posterior using a Bayesian updating table and doing the necessary algebra.

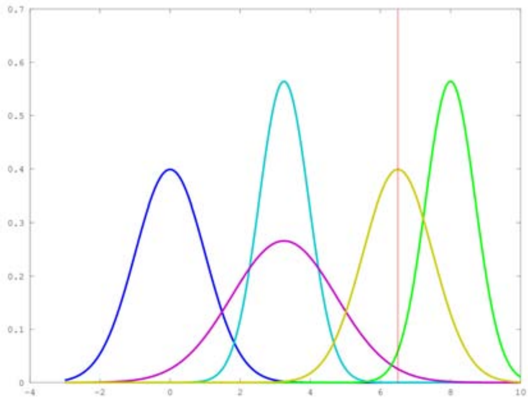
3. Understand that the updating formulas come by using the updating tables and doing the algebra.

Concept question

$X \sim \text{N}(\theta, \sigma^2)$; $\sigma = 1$ is known.

Prior pdf at far left in blue; single data point marked with red line.

Which is the posterior pdf?



1. Cyan
2. Magenta
3. Yellow
4. Green

Conjugate priors

Priors pairs that update to the same type of distribution.
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

Board question: normal/normal

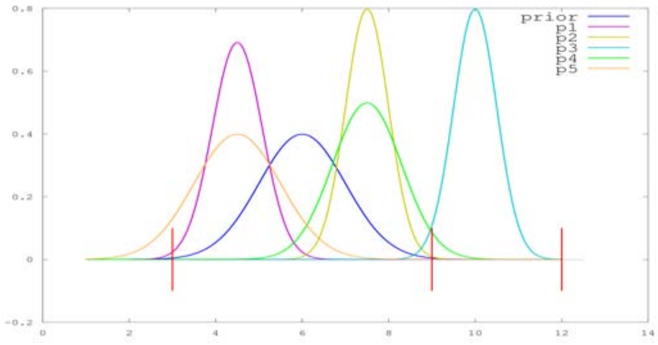
For data $x_1, \ldots, x_n$ with data mean $\bar{x} = \frac{x_1 + \ldots + x_n}{n}$

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

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

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

Concept question: normal priors, normal likelihood

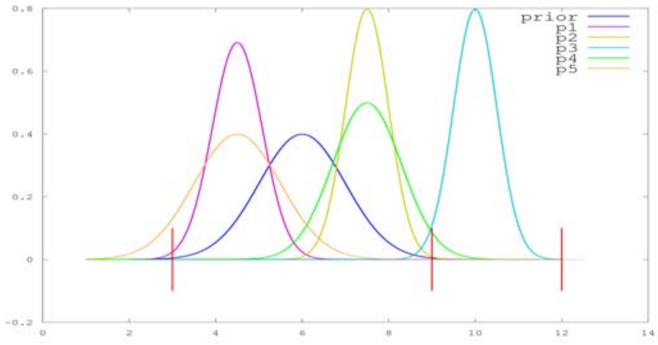


Blue = prior   Red = data in order: 3, 9, 12

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

1. blue
2. magenta
3. orange
4. yellow
5. green
6. light blue

Concept question: normal priors, normal likelihood



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

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

1. blue
2. magenta
3. orange
4. yellow
5. green
6. light blue

Variance can increase

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

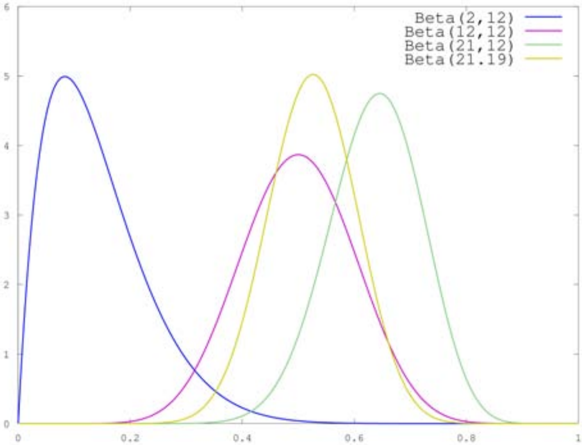


Table discussion: likelihood principle

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

(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 the 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.

The true statements are:

1. all true	**2**. a,b,c	**3**. a,b,d	**4**. a,c	**5**. d.

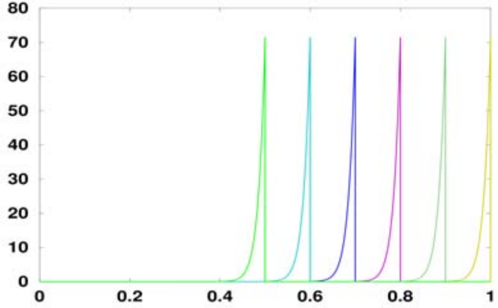
Concept question

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

We are convinced that $\theta \leq .7$.
Our prior is uniform on [0,.7] and 0 from .7 to 1.

We flip the coin 65 times and get 60 heads.

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



1. green	**2.** light blue	**3.** blue
4. magenta	**5.** light green	**6.** yellow