Bayesian Updating: Continuous Priors

18.05 Spring 2014

Jeremy Orloff and Jonathan Bloom

Beta distribution

$Beta(a, b)$ has density

$$f(\theta) = \frac{(a+b-1)!}{(a-1)!(b-1)!}\theta^{a-1}(1-\theta)^{b-1}$$

http://ocw.mit.edu/ans7870/18/18.05/s14/applets/beta-jmo.html

Observation:

The coefficient is a normalizing factor, so if we have a pdf

$$f(\theta) = c\theta^{a-1}(1-\theta)^{b-1}$$

then

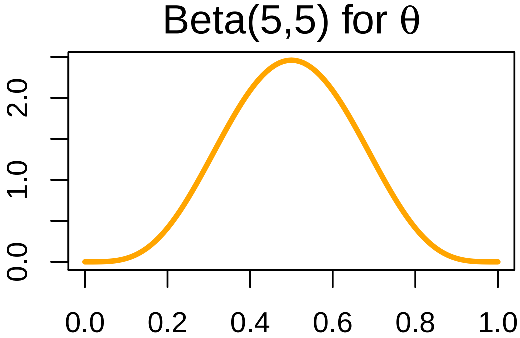
$$\theta \sim \text{beta}(a, b)$$

and

$$c = \frac{(a+b-1)!}{(a-1)!(b-1)!}$$

Board question preamble: beta priors

Suppose you have a coin with unknown probability of heads θ. You don't know that it's fair, but your prior belief is that it's probably not too unfair. You capture this intuition in with a beta(5,5) prior on θ.



In order to sharpen this distribution you take data and update the prior.

Question on next slide.

Board question: beta priors

- $Beta(a,b)$: $f(\theta) = \dfrac{(a+b-1)!}{(a-1)!(b-1)!}\theta^{a-1}(1-\theta)^{b-1}$
- Coin has prior $f(\theta) \sim \text{beta}(5,5)$

1. Suppose you flip 10 times and get 6 heads. Find the posterior distribution on θ. Identify the type of the posterior distribution.

2. Suppose you recorded the order of the flips and got H H H T T H H H T T. Find the posterior based on this data.

3. Using your answer to (2) give an integral for the posterior predictive probability of heads on the next toss.

4. Use what you know about pdf's to evaluate the integral without computing it directly

Predictive probabilities

Continuous hypotheses θ, discrete data $x_1, x_2, \ldots$
(Assume trials are independent.)

Prior predictive probability

$$p(x_1) = \int p(x_1 \,|\, \theta) f(\theta)\, d\theta$$

Posterior predictive probability

$$p(x_2 \,|\, x_1) = \int p(x_2 \,|\, \theta) f(\theta \,|\, x_1)\, d\theta$$

Analogous to discrete hypotheses: $\mathcal{H}_1, \mathcal{H}_2, \ldots$.

$$p(x_1) = \sum_{i=1}^{n} p(x_1 \,|\, \mathcal{H}_i) P(\mathcal{H}_i) \qquad p(x_2 \,|\, x_1) = \sum_{i=1}^{n} p(x_2 \,|\, \mathcal{H}_i) p(\mathcal{H}_i \,|\, x_1).$$

Concept Question

Suppose your prior $f(\theta)$ in the bent coin example is Beta$(6,8)$. You flip the coin 7 times, getting 2 heads and 5 tails. What is the posterior pdf $f(\theta|x)$?

1. Beta(2,5)

2. Beta(3,6)

3. Beta(6,8)

4. Beta(8,13)

Continuous priors, continuous data

Bayesian update tables with and without infinitesimals

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

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

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

Normal prior, normal data

$\mathrm{N}(\mu, \sigma^2)$ has density

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-(y-\mu)^2/2\sigma^2}.$$

Observation:

The coefficient is a normalizing factor, so if we have a pdf

$$f(y) = c\mathrm{e}^{-(y-\mu)^2/2\sigma^2}$$

then

$$y \sim \mathrm{N}(\mu, \sigma^2)$$

and

$$c = \frac{1}{\sigma\sqrt{2\pi}}$$

Board question: normal prior, normal data

- $\mathrm{N}(\mu, \sigma^2)$ has pdf: $\quad f(y) = \dfrac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-(y-\mu)^2/2\sigma^2}$.
- Suppose our data follows a $\mathrm{N}(\theta, 4)$ distribution with unknown mean θ and variance 4. That is

$$f(x \,|\, \theta) = \text{pdf of } \mathrm{N}(\theta, 4)$$

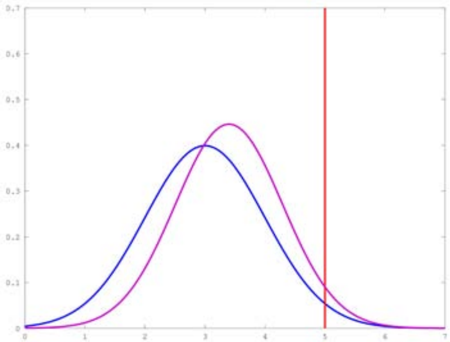
- Suppose our prior on θ is $\mathrm{N}(3, 1)$.

Suppose we obtain data $x_1 = 5$.

1. Use the data to find the posterior pdf for θ.

Write out your tables clearly. Use (and understand) infinitesimals.

Solution graphs



prior = blue; posterior = purple; data = red

Data:	$x_1 = 5$	
Prior:	$\mu_{\text{prior}} = 3;$	$\sigma_{\text{prior}} = 1$
Posterior is normal	$\mu_{\text{posterior}} = 3.4;$	$\sigma_{\text{posterior}} = 0.894$

Board question: Romeo and Juliet

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

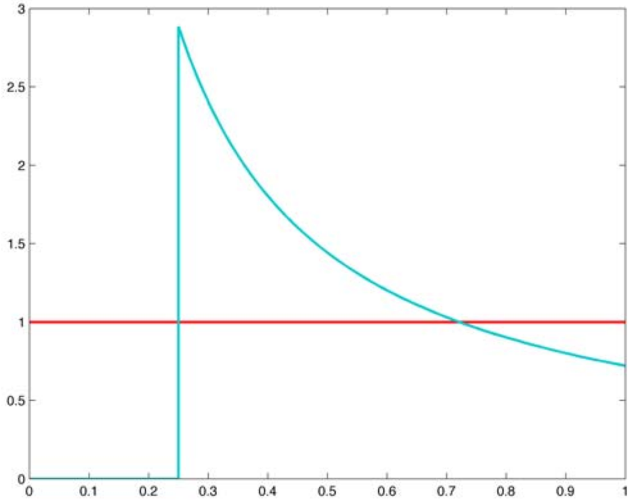
Juliet knows that $\theta \leq 1$ hour and she assumes a flat prior for θ on $[0, 1]$.

On their first date Romeo is 15 minutes late.

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

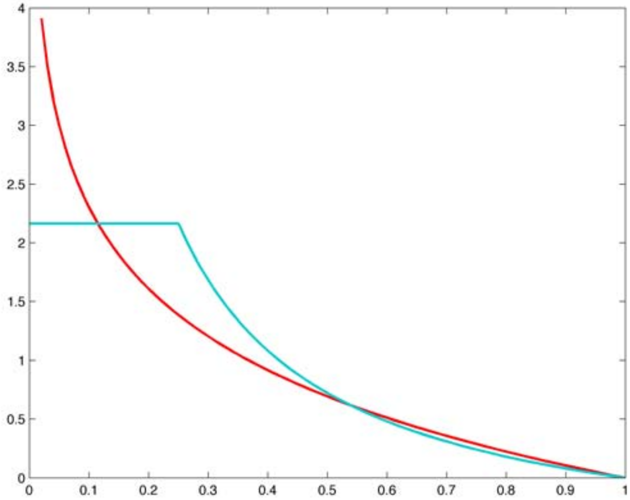
(b) find and graph the prior predictive and posterior predictive pdf's of how late Romeo will be on the second data (if he gets one!).

Solution continued



Prior and posterior pdf's for θ.

Solution continued



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

From discrete to continuous Bayesian updating

Bent coin with unknown probability of heads θ.

Data x_1: heads on one toss.

Start with a flat prior and update:

hyp.	prior	likelihood	unnormalized posterior	posterior
θ	$d\theta$	θ	$\theta\, d\theta$	$2\theta\, d\theta$
Total	1		$\int_0^1 \theta\, d\theta = 1/2$	1

Posterior pdf: $f(\theta \,|\, x_1) = 2\theta.$

Approximate continuous by discrete

- approximate the continuous range of hypotheses by a finite number of hypotheses.
- create the discrete updating table for the finite number of hypotheses.
- consider how the table changes as the number of hypotheses goes to infinity.

Chop $[0, 1]$ into 4 intervals

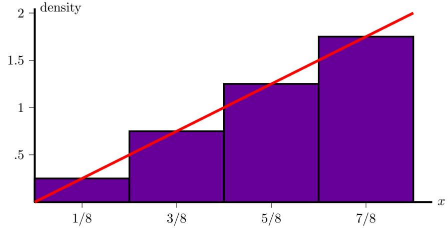
hypothesis	prior	likelihood	un. posterior	posterior
$\theta = 1/8$	$1/4$	$1/8$	$(1/4) \times (1/8)$	$1/16$
$\theta = 3/8$	$1/4$	$3/8$	$(1/4) \times (3/8)$	$3/16$
$\theta = 5/8$	$1/4$	$5/8$	$(1/4) \times (5/8)$	$5/16$
$\theta = 7/8$	$1/4$	$7/8$	$(1/4) \times (7/8)$	$7/16$
Total	1	–	$\sum_{i=1}^{n} \theta_i \, \Delta\theta$	1

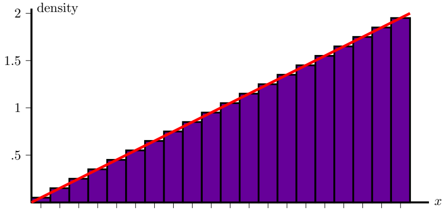
Chop $[0, 1]$ into 12 intervals

hypothesis	prior	likelihood	un. posterior	posterior
$\theta = 1/24$	$1/12$	$1/24$	$(1/12) \times (1/24)$	$1/144$
$\theta = 3/24$	$1/12$	$3/24$	$(1/12) \times (3/24)$	$3/144$
$\theta = 5/24$	$1/12$	$5/24$	$(1/12) \times (5/24)$	$5/144$
$\theta = 7/24$	$1/12$	$7/24$	$(1/12) \times (7/24)$	$7/144$
$\theta = 9/24$	$1/12$	$9/24$	$(1/12) \times (9/24)$	$9/144$
$\theta = 11/24$	$1/12$	$11/24$	$(1/12) \times (11/24)$	$11/144$
$\theta = 13/24$	$1/12$	$13/24$	$(1/12) \times (13/24)$	$13/144$
$\theta = 15/24$	$1/12$	$15/24$	$(1/12) \times (15/24)$	$15/144$
$\theta = 17/24$	$1/12$	$17/24$	$(1/12) \times (17/24)$	$17/144$
$\theta = 19/24$	$1/12$	$19/24$	$(1/12) \times (19/24)$	$19/144$
$\theta = 21/24$	$1/12$	$21/24$	$(1/12) \times (21/24)$	$21/144$
$\theta = 23/24$	$1/12$	$23/24$	$(1/12) \times (23/24)$	$23/144$
Total	1	–	$\sum_{i=1}^{n} \theta_i \, \Delta\theta$	1

Density historgram

Density historgram for posterior pmf with 4 and 20 slices.





The original posterior pdf is shown in red.