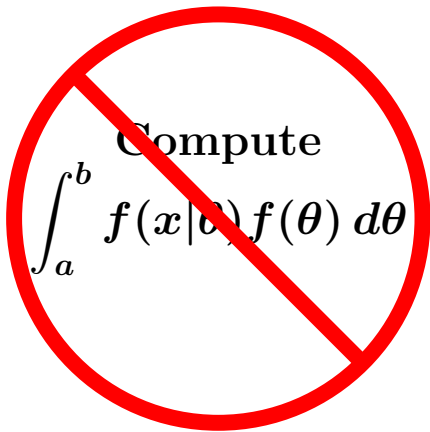


Bayesian Updating: Continuous Priors

18.05 Spring 2014



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://mathlets.org/mathlets/beta-distribution/>

Observation:

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

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

then

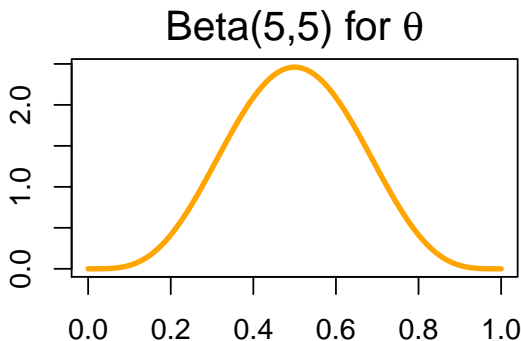
$$\theta \sim \text{beta}(a, b)$$

and

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

Board question preamble: beta priors

Suppose you are testing a new medical treatment with unknown probability of success θ . You don't know that θ , but your prior belief is that it's probably not too far from 0.5. You capture this intuition with a beta(5,5) prior on θ .



To sharpen this distribution you take data and update the prior.

Question on next slide.

Board question: beta priors

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

1. Suppose you test it on 10 patients and have 6 successes. Find the posterior distribution on θ . Identify the type of the posterior distribution.
2. Suppose you recorded the order of the results and got SSSFFSSSFF. Find the posterior based on this data.
3. Using your answer to (2) give an integral for the posterior predictive probability of success with the next patient.
4. Use what you know about pdf's to evaluate the integral without computing it directly

Conjugate priors

We had

- Prior $f(\theta) d\theta$: **beta distribution**
- Likelihood $p(x|\theta)$: binomial distribution
- Posterior $f(\theta|x) d\theta$: **beta distribution**

The beta distribution is called a **conjugate prior** for the binomial likelihood.

That is, the beta prior becomes a beta posterior and repeated updating is easy!

Concept Question

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

1. $\text{Beta}(2,5)$
2. $\text{Beta}(3,6)$
3. $\text{Beta}(6,8)$
4. $\text{Beta}(8,13)$

Reminder: predictive probabilities

Continuous hypotheses θ , discrete data x_1, x_2, \dots

(Assume trials are independent given the hypothesis θ .)

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: $\mathcal{H}_1, \mathcal{H}_2, \dots$

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

Continuous priors, continuous data

Bayesian update tables:

hypoth.	prior	likelihood	Bayes numerator	posterior $f(\theta x) d\theta$
θ	$f(\theta) d\theta$	$f(x \theta)$	$f(x \theta)f(\theta) d\theta$	$\frac{f(x \theta)f(\theta) d\theta}{f(x)}$
total	1		$f(x)$	1

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

Normal prior, normal data

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

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

Observation:

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

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

then

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

and

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

Board question: normal prior, normal data

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

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

- Suppose our prior on θ is $N(3, 1)$.

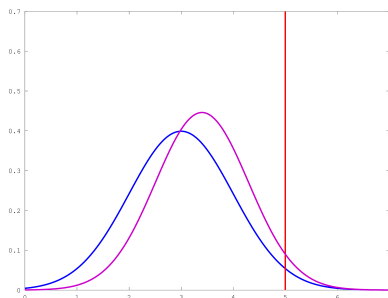
Suppose we obtain data $x_1 = 5$.

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

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

You will have to remember how to complete the square to do the updating!

Solution graphs



prior = blue; posterior = purple; data = red

Data: $x_1 = 5$

Prior is normal: $\mu_{\text{prior}} = 3$; $\sigma_{\text{prior}} = 1$

Likelihood is normal: $\mu = \theta$; $\sigma = 2$

Posterior is normal $\mu_{\text{posterior}} = 3.4$; $\sigma_{\text{posterior}} = 0.894$

- Will see simple formulas for doing this update next time.

Board question: Romeo and Juliet

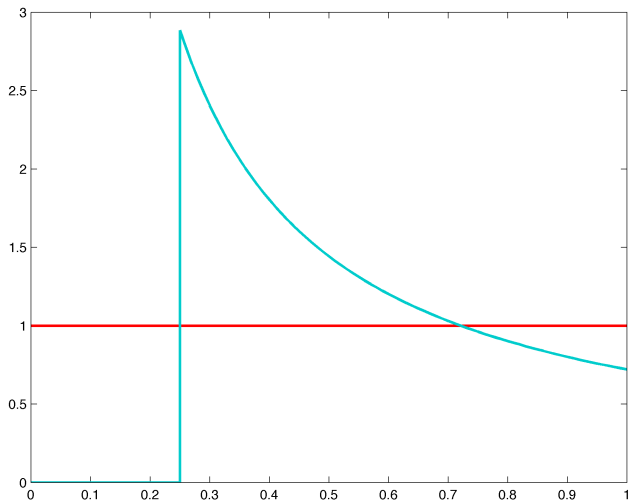
Romeo is always late. How late follows a uniform distribution $\text{uniform}(0, \theta)$ with unknown parameter θ in hours.

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. Use this data to update the prior distribution for θ .

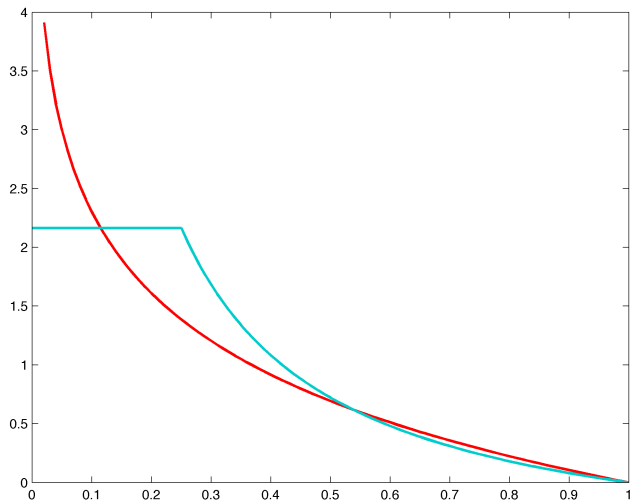
- (a) Find and graph the prior and posterior pdfs for θ .
- (b) Find the prior predictive pdf for how late Romeo will be on the first date and the posterior predictive pdf of how late he'll be on the second date (if he gets one!). Graph these pdfs.

Solution graphs



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	Bayes numerator	numerator
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

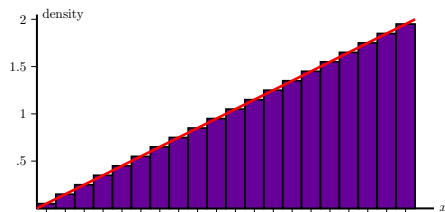
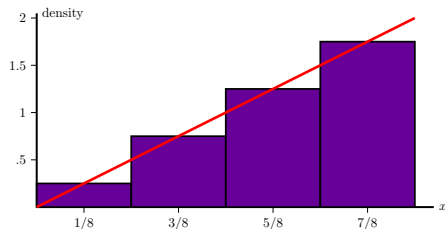
hypothesis	prior	likelihood	Bayes num.	posterior
$\theta = 1/8$	$1/4$	$1/8$	$(1/4) \times (1/8)$	$1/16$
$\theta = 3/8$	$1/4$	$3/8$	$(1/4) \times (3/8)$	$3/16$
$\theta = 5/8$	$1/4$	$5/8$	$(1/4) \times (5/8)$	$5/16$
$\theta = 7/8$	$1/4$	$7/8$	$(1/4) \times (7/8)$	$7/16$
Total	1	-	$\sum_{i=1}^n \theta_i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

Chop $[0, 1]$ into 12 intervals

hypothesis	prior	likelihood	Bayes num.	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 histogram

Density histogram for posterior pmf with 4 and 20 slices.



The original posterior pdf is shown in red.

MIT OpenCourseWare
<https://ocw.mit.edu>

18.05 Introduction to Probability and Statistics

Spring 2014

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.