Studio 5: Bayesian Updating
18.05 Spring 2014
Jeremy Orloff and Jonathan Bloom

You should have downloaded studio5.zip and unzipped it into your 18.05 working directory.

Board question

1. Consider the random die problem: 5 dice with sides 4, 6, 8, 12 or 20. I choose one die at random and roll it repeatedly, reporting the results to you. Your job is to figure out which die I chose.

(a) List all possible hypotheses.

(b) List all possible outcomes for the jth roll.

(c) Let x_j be the random variable equal to the outcome of the jth roll, $j = 1, 2, \ldots$.
Write out the entire likelihood table for the jth roll. The table has a column of likelihoods for each possible outcome.

Solutions on next slide.

Solution

(a) The hypotheses are that the die has 4, 6, 8, 12, or 20 sides. Let's call these H_4, H_6 etc.

(b) The possible outcomes are the numbers 1 to 20.

(c) The rows in the table are hypotheses; the columns are outcomes. To save space, we combine outcomes that give the same likelihoods into one column.

x_j is the outcome of the jth roll.

Hypothesis	$x_j = 1\text{-}4$	$x_j = 5\text{-}6$	$x_j = 7\text{-}8$	$x_j = 9\text{-}12$	$x_j = 13\text{-}20$
H_4	1/4	0	0	0	0
H_6	1/6	1/6	0	0	0
H_8	1/8	1/8	1/8	0	0
H_{12}	1/12	1/12	1/12	1/12	0
H_{20}	1/20	1/20	1/20	1/20	1/20

R question

2. (a) Go through the simulation of Baysian updating in studio5.r. After running through the loop of 20 rolls be sure to click through all the plots in RStudio.

(b) Use the same `dataRolls` as in part (a) and redo the simulation with a new prior distribution.

$\mathcal{H}$	4	6	8	12	20
$P(\mathcal{H})$	.05	.05	.05	.05	.80

(c) Describe how the posterior distributions compare.

(d) Redo the simulation with the original prior distribution but generate the sample rolls `dataRolls` from a 20-sided die.

Censored data

3. Suppose we only observe **censored** values of the rolls. For the jth roll we observe

$$y_j = \begin{cases} 1 & if & x_j = 1 \\ 0 & if & x_j \neq 1 \end{cases}$$

Instead of the actual rolls you only see the censored data $y_1, y_2, \ldots$.
(a) List all the possible hypotheses.

(b) Write out the entire likelihood table for one roll y_j.

(c) Modify the R simulation to simulate this censored scenario.

- Use nrolls = 20 in the simulation. After the 20th roll has the posterior nearly converged on the true hypothesis?
- Investigate how many rolls are needed to see a clear convergence of the posterior distribution.

18.05 Introduction to Probability and Statistics

Spring 2014