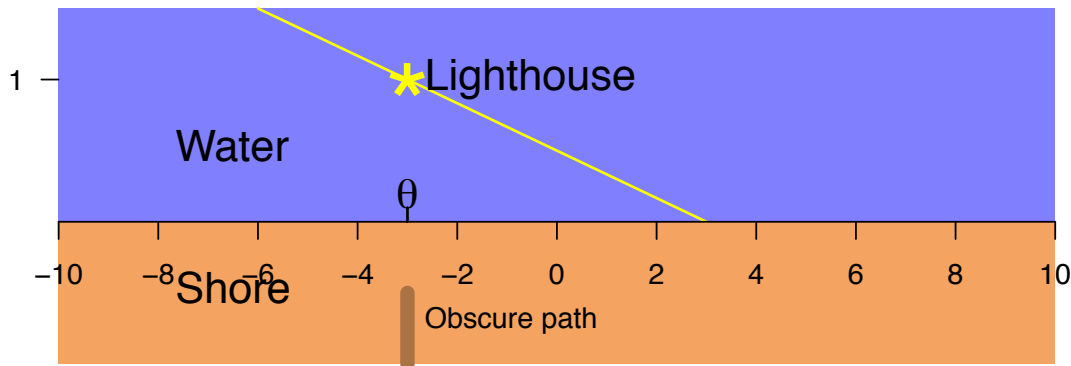


Studio 6 Discretized Continuous Bayesian Updating

18.05, Spring 2022

2. Shore-sea-lighthouse



Overview of the studio

This studio explores Bayesian updating of continuous priors.

R introduced in this studio

The R needed is introduced in `mit18_05_s22_studio6-samplecode.r`.

New functions: `dcauchy()`, `which.max()`.

Concepts: Bayesian updating, discretization, Maximum a posteriori (MAP) estimates, Cauchy distribution.

Download the zip file

- You should have downloaded the file `mit18_05_s22_studio6.zip` from our MITx site.
- Unzip it in your 18.05 studio folder.
- It contains the following R files that you will work with
`mit18_05_s22_studio6.r`, `mit18_05_s22_studio6-samplecode.r`,
`mit18_05_s22_studio6-test.r`
- The following files will help you write and test your code
`mit18_05_s22_studio6-test-answers.html`
- It also has the following files that that your code will run. You will not need to look at these. (Although it might be interesting)
`mit18_05_s22_studio6_problem0_draw_plot.r`, `mit18_05_s22_studio6_problem_1d.r`,
`mit18_05_s22_studio6_data_frame.csv`

Prepping R Studio

- In R studio, open `mit18_05_s22_studio6-samplecode.r` and `mit18_05_s22_studio6.r`
- Using the Session menu, set the working directory to source file location. (This is a good habit to develop!)
- Answer the questions in the detailed instructions just below. Your answers should be put in `mit18_05_s22_studio6.r`
- [Solution code will be posted tomorrow at 10 pm](#)

Detailed instructions for the studio

- Go through `mit18_05_s22_studio6-samplecode.r` as a tutorial.

Problem 0

This problem only requires a short description of what you see.

- Run the function `studio6_problem0()`.

This will run the source file `studio6_problem0_draw_plot.r`, which creates two histograms:

Histogram 1: (Simulated) data from a normal distribution $\text{Norm}(10, 6^2)$.

Histogram 2: (Simulated) averages of 9 draws from $\text{Norm}(10, 6^2)$.

- In the cat statement at the end of this function, describe the effect of averaging on the distribution.

Problem 1

This problem introduces the Cauchy distribution.

Problem 1a. Look up the Cauchy distribution on Wikipedia.

In the cat statement at the end of the function `studio6_problem_1a()`, give the formula for the Cauchy pdf with location `theta` and `scale = 1.0`.

Problem 1b. Here you will finish the code for the function `studio6_problem_1b()`.

For this function you should add code that

- On one set of axes, plots the pdfs of standard normal and the Cauchy distribution with `location = 0` and `scale = 1.0`.
- Plot the Cauchy in blue and the normal in orange.

Problem 1c. Here you will complete the cat statement for the function `studio6_problem_1c()`.

Looking at the graphs drawn in 1b, explain why we say the Cauchy distribution has fat tails when compared with the standard normal.

Problem 1d. Just call the function `studio6_problem1d()`.

This function mirrors problem 0, except it uses the Cauchy distribution.

There is nothing for you to do here, but admire the amazing variability of the natural world. Notice that the histogram of the average data does not get narrower. This is a feature of the Cauchy distribution – remarkably, the distribution of the average of i.i.d. Cauchy values follows the exact same Cauchy distribution.

Problem 2: Cauchy’s Lighthouse

The sample code in `mit18_05_s22_studio6-samplecode.r`, will be quite helpful for this problem.

Here is the setting for this problem.

A lighthouse stands one kilometer offshore. Its bi-directional beam rotates at a constant rate, so that at any moment there is exactly one spot on the straight shoreline illuminated by the beam. All along the shore are kilometer markers.

A picture of the scene is given on page one of this file.

You arrived during the day and didn’t note the location of the path. Now it is now night and to get home you need to find an obscure path near the shore at the same kilometer marker as the lighthouse. Your only hope for finding the path is to get a good estimate of the position of the lighthouse. As you wander up and down the shore, occasionally you and the light from the lighthouse coincide and you can collect the datum of the position where the beam hits the shore. You manage to do this 15 times.

We can show (and you may assume) the position data follows a Cauchy distribution with unknown location = θ and scale = 1.0.

In this problem we will use Bayesian updating to make an estimate of the value of θ .

Problem 2a. Here you will finish the code for the function: `studio6_problem_2a(data_frame_csv)`

Arguments:

`data_frame_csv` = name of the data file. (The grader will use a different data file than the test scripts.)

The position data you collected is stored in the file `mit18_05_s22_studio6_data_frame.csv`. The code we give you at the start of the function loads the data . (We will use different data when we grade your code.)

It’s usually a good idea to just plot data and see if anything jumps out at you. Add code that plots the position data. (See `mit18_05_s22_studio6-samplecode.r` section 0.)

Problem 2b. In this problem you will finish code for the function `mit18_05_s22_studio6_problem_2b(data_frame_csv)`.

Arguments:

`data_frame_csv` = name of the data file. (The grader will use a different data file than the test scripts.)

This is the one long problem in this assignment. Sections 1-4 of the sample code will be

helpful.

Assume a $\text{Uniform}(\theta_{\min}, \theta_{\max})$ prior for θ . In the code given at the start of the function we assign values `theta_min` and `theta_max` for θ_{\min} and θ_{\max} .

You need to write code to do the following.

2(i) Discretize the prior: use the value `dtheta` given in the function as the discretization stepsize.

2(ii) Use a for loop to go through the data one value at a time and do Bayesian updating. Save each posterior as you go. (Before the loop, you will need to initialize a matrix to hold the posteriors.)

2(iii) On one plot graph the prior and each of the posteriors

2(iv) The maximum a posteriori (MAP) estimate is the value of θ that maximizes the posterior function. Find the MAP estimate for each of the posteriors. Save these in an array and plot them.

2(v) Make a separate plot showing the final posterior pdf and its MAP estimate.

2(vi) In the cat statement at the end of the function, say where, given the final posterior, you would look for the obscure path.

OPTIONAL Problem 2c. This problem is optional. In it you will have `studio6_problem_2c()` print out an explanation for why the data follows a Cauchy distribution.

Testing your code

For each problem, we ran the problem function with certain parameters. You can see the function call and the output in `mit18_05_s22_studio6-test-answers.html`. If you call the same function with the same parameters, you should get the same results as in `mit18_05_s22_studio6-test-answers.html` – if there is randomness involved the answers should be close but not identical.

For your convenience, the file `mit18_05_s22_studio6-test.r` contains all the function calls used to make `mit18_05_s22_studio6-test-answers.html`.

Before uploading your code

1. Make sure all your code is in `mit18_05_s22_studio6.r`. Also make sure it is all inside the functions for the problems.
2. Clean the environment and plots window.
3. Source the file.
4. Call each of the problem functions with the same parameters as the test file `mit18_05_s22_studio6-test-answers.html`.
5. Make sure it runs without error and outputs just the answers asked for in the questions.
6. Compare the output to the answers given in `mit18_05_s22_studio6-test-answers.html`.

Upload your code

Use the upload link on our MITx site to upload your code for grading.

Leave the file name as `mit18_05_s22_studio6.r`. (The upload script will automatically add your name and a timestamp to the file.)

You can upload more than once. We will grade the last file you upload.

Due date

Due date: The goal is to upload your work by the end of class.

If you need extra time, you can upload your work any time before 10 PM ET the day after the studio.

Solutions uploaded: Solution code will be posted on MITx at 10 PM the day after the studio.

MIT OpenCourseWare

<https://ocw.mit.edu>

18.05 Introduction to Probability and Statistics

Spring 2022

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