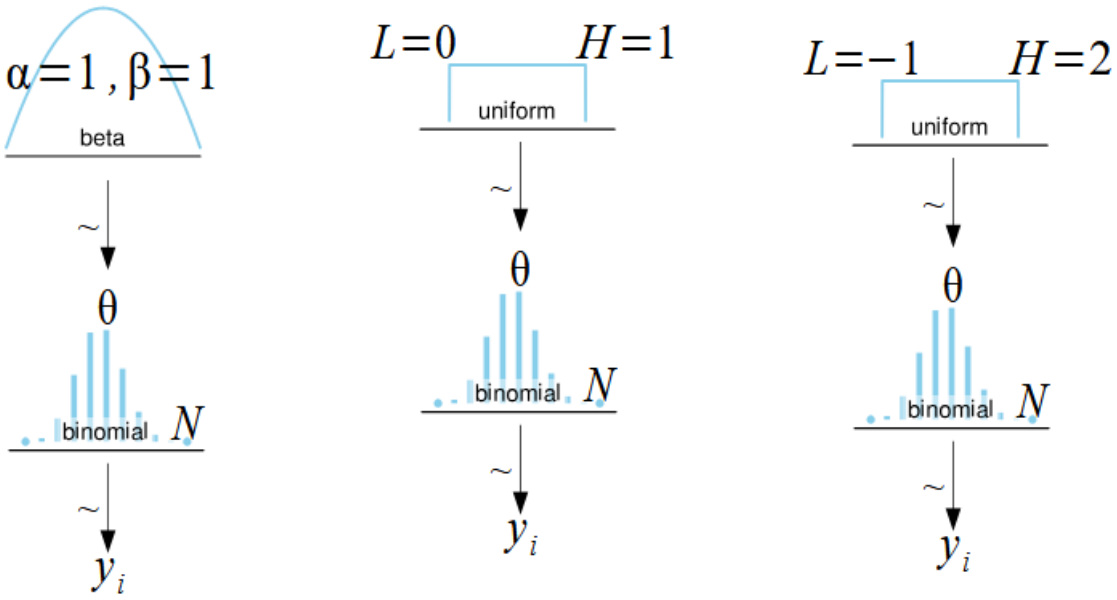
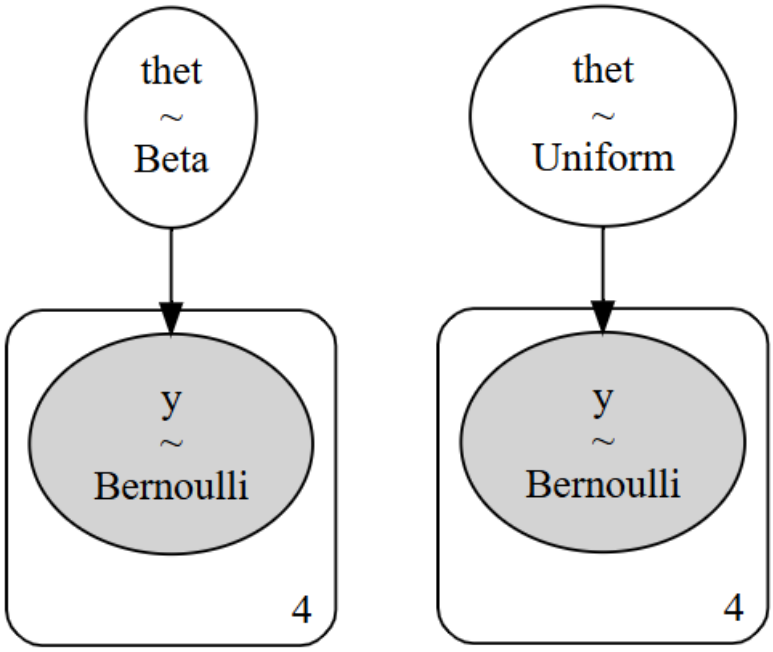
Homework 2 Solutions

Question 1:

a.

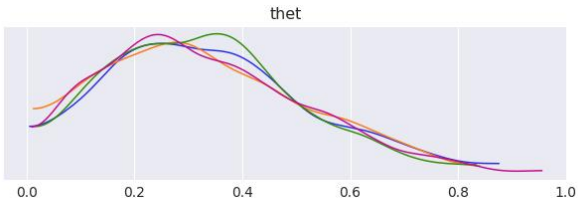


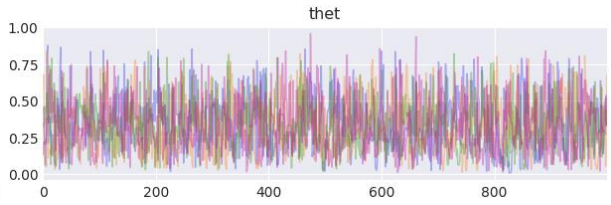
b.

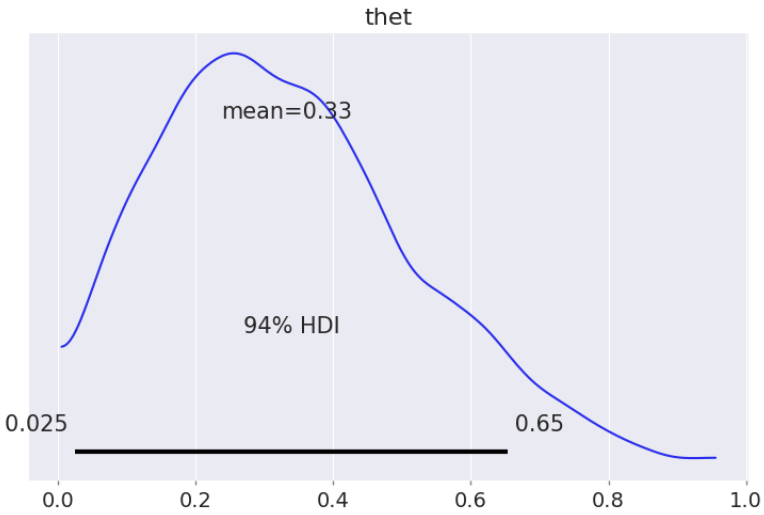


c. The model with the uniform distribution (-1, 2) runs but has many divergences.

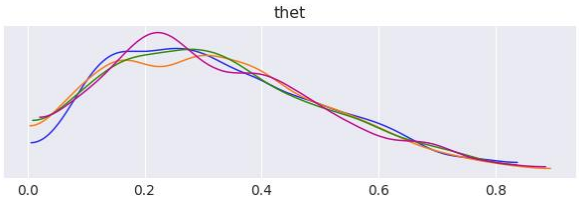
d. Beta prior:

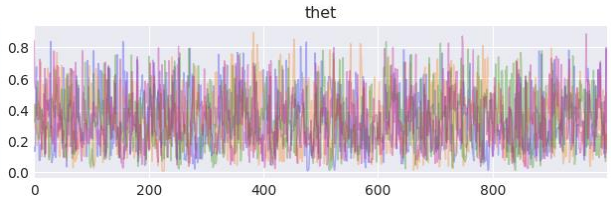


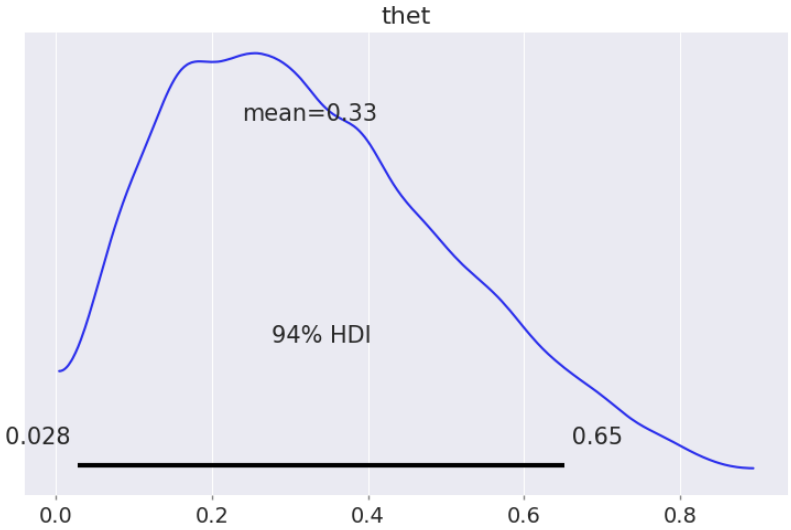




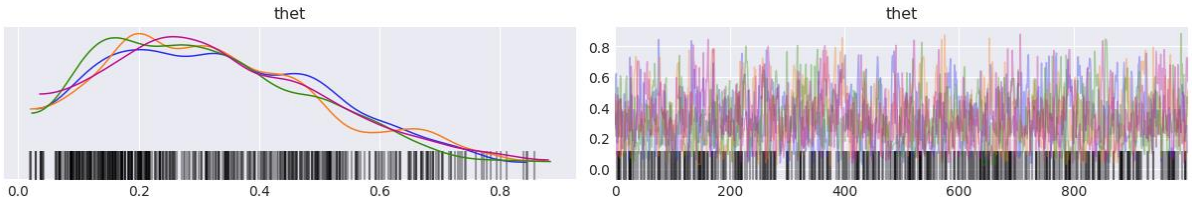
Uniform prior [0, 1]:

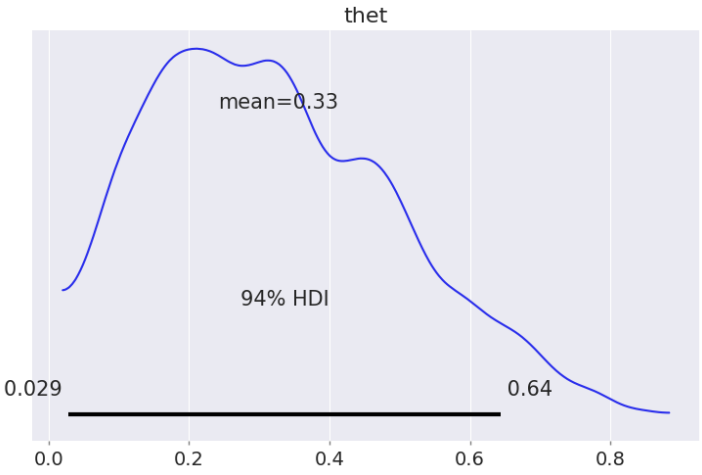






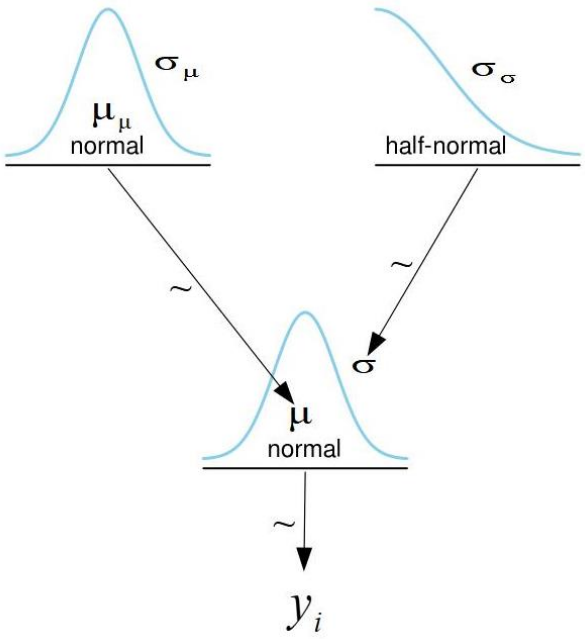
Uniform prior [-1, 2]:





Question 2:

b.



$$y_i \sim N(\mu, \sigma)$$

$$\mu \sim N(\mu_\mu, \sigma_\mu)$$

$$\sigma \sim HalfNorm(\sigma_\sigma)$$

c. We chose standard deviation values equal to that of the data (~3.5), three times that of the data (~10) and 10 times that of the data (~35).

See notebook for results.

d. We can see that it worked relatively well. As long as we set priors on our μ and σ that together with the data and sampling, keep the posterior in the range that makes sense, then it can work well. Normal models also have the advantage of running faster. The disadvantage is that we can get values that are impossible. This will also be affected by our data. For example, if we have mainly values that are far from zero, there's less of a chance of wasting data keeping the model in the right bounds. Another disadvantage is the lack of clear communication. Using a bounded distribution clearly communicates to someone else looking at the model that our values are bounded.

Question 3:

See notebook for results.

Question 4:

a.
```
Removing outliers removed 2 data points
The mean and standard devation of all the data is (np.float64(53.5), np.float64(3.46))
The mean and standard devation of the data without outliers is (np.float64(52.95), np.float64(2.22))
```

b. We will look at the posterior distributions for the mean and standard deviation. Then we'll look at the mean of those distributions.

For the normal model we got: mean = 54, std = 3.5.

For the t-distribution we got: mean = 53, std = 2.1.

We can see from this that the normal distribution's standard deviation is wider to account for the outliers. On the other hand, the standard deviation of the t-distribution is narrower, and very similar to that of the data without the outliers. Hence, the t-distribution succeeds in capturing the standard deviation of the bulk of the data, as it can use it's normality parameter (defining the kurtosis) to account for the outliers in the tails, and does not need to be wider to account for this.

c. See notebook.

Question 5:

b. 1. 4 for four days.

2. See notebook.

3. See notebook for plot.

The dimensions are sample, and the coordinates are the sample, chain and draw.