Homework 6

David L. Olson

May 8, 2025

Contents

1 Problem 2

Exercise 10 in Section 11.6

1.1 Solution

1.1.1 Part A

The first step is to convert the grid of χ^2 values from the previous homework problem into the likelihood values $L(\boldsymbol{m}|\boldsymbol{d})$. Recall the surface of chi^2 values from the previous problem below.

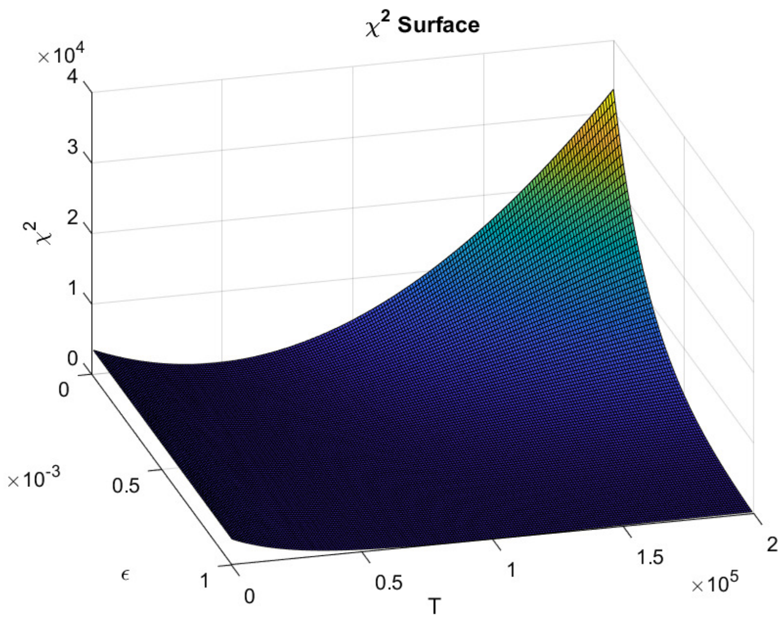

Figure 1: Chi-Squared Surface

The likelihood is equal to $L(\boldsymbol{m}|\boldsymbol{d}) = f(\boldsymbol{d}|\boldsymbol{m}) = f(d_1|\boldsymbol{m})\, f(d_1|\boldsymbol{m}) \ldots f(d_m|\boldsymbol{m})$.
The likelihood of a given data point d_i is below.

$$f(d_i|\boldsymbol{m}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(G(\boldsymbol{m})_i - d_i)^2}{2\sigma^2}}$$

For m data points, the product of each $f(d_i|\boldsymbol{m})$ is also shown below, which differs slightly from equation 11.10 since our data points in this problem have different sigma values.

$$L(\boldsymbol{m}|\boldsymbol{d}) = f(\boldsymbol{d}|\boldsymbol{m}) = \left[\prod_{i=1}^{m} \frac{1}{\sigma_i\sqrt{2\pi}}\right] \left[e^{-\sum_{i=1}^{m} \frac{(G(\boldsymbol{m})_i - d_i)^2}{2\sigma^2}}\right]$$

Recall that $\chi^2 = \sum_{i=1}^{m} \frac{(G(\boldsymbol{m})_i - d_i)^2}{\sigma^2}$ per equation 2.20, which can be substituted in the above expression.

$$L(\boldsymbol{m}|\boldsymbol{d}) = \left[\prod_{i=1}^{m} \frac{1}{\sigma_i\sqrt{2\pi}}\right]\left[e^{\frac{-\chi^2}{2}}\right]$$

Doing so allows us to convert our grid of chi^2 values into likelihood values. Given that the grid consists of so many points, the likelihood at each point in the grid is incredibly small.

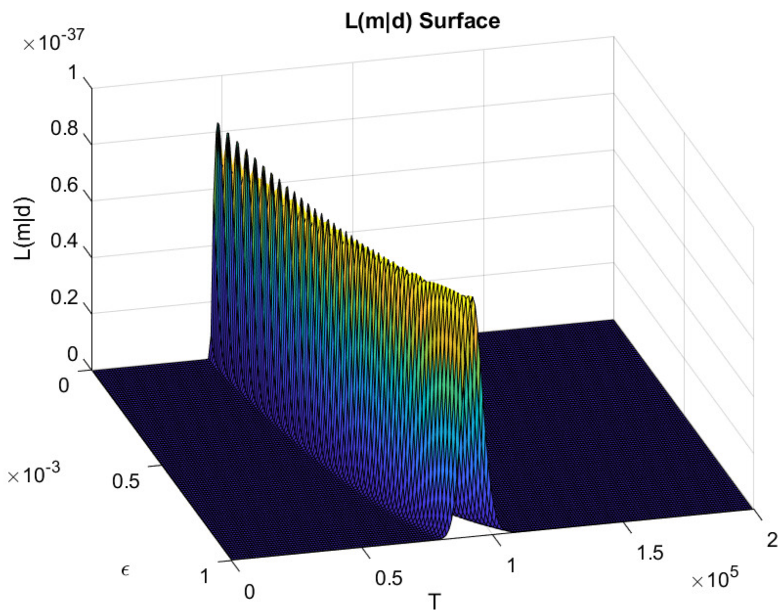

Figure 2: Likelihood Surface

The posterior distribution is given by the expression below:

$$q(\boldsymbol{m}|\boldsymbol{d}) = \frac{p(\boldsymbol{m})\, f(\boldsymbol{d}|\boldsymbol{m})}{\int_{\text{all models}} f(\boldsymbol{d}|\boldsymbol{m})\, p(\boldsymbol{m})\, d\boldsymbol{m}}$$

The prior is a uniform distribution $P(\epsilon, T) = 1$ which simplifies the above expression as shown below.

$$q(\boldsymbol{m}|\boldsymbol{d}) = \frac{f(\boldsymbol{d}|\boldsymbol{m})}{\int_{\text{all models}} f(\boldsymbol{d}|\boldsymbol{m})\, d\boldsymbol{m}}$$

In essence, the posterior distribution is $f(\boldsymbol{d}|\boldsymbol{m})$ normalized by the volume under the two-dimensional surface. The volume can be approximated as:

$$q\left(\boldsymbol{m}|\boldsymbol{d}\right) \approx \frac{f\left(\boldsymbol{d}|\boldsymbol{m}\right)}{\sum f\left(\boldsymbol{d}|\boldsymbol{m}\right) d\epsilon dT}$$

The resulting posterior distribution is shown in figure 3;

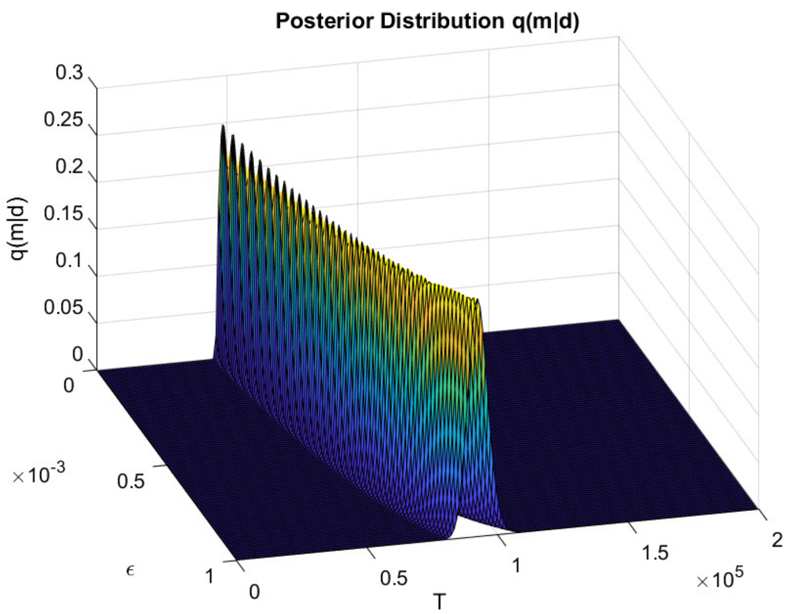

Figure 3: Posterior Distribution

1.1.2 Part B

Given the joint probability $q\left(\boldsymbol{m}|\boldsymbol{d}\right) = p\left(\epsilon, T\right)$, each marginal probability for the two random variables are below.

$$p_\epsilon(\epsilon) = \int_T p\left(\epsilon, T\right) dT$$

$$p_T(T) = \int_\epsilon p\left(\epsilon, T\right) d\epsilon$$

The marginal probability density functions (technically, its a probability mass function) are in figure 4.

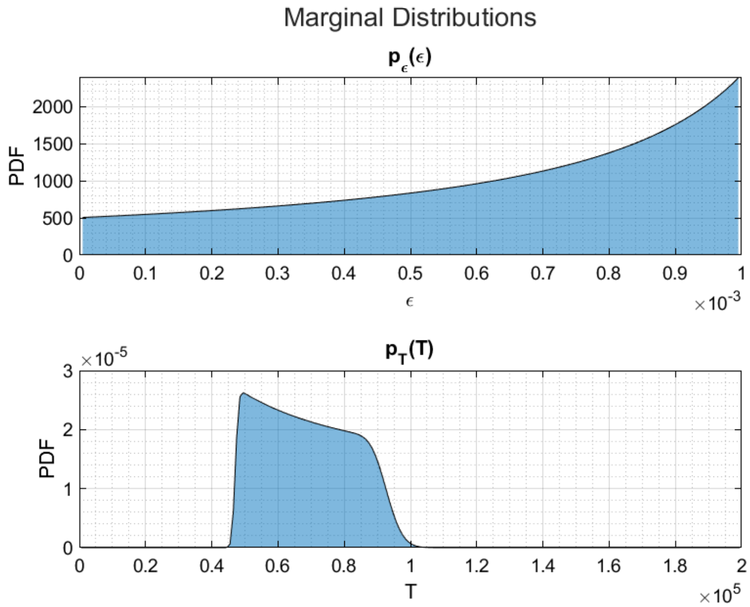

Figure 4: Marginal Distributions

1.1.3 Part C

Now we are provided a new prior, which is shown in figure 5.

$$p(\boldsymbol{m}) = p(\epsilon, T) \propto e^{-\frac{\left(\epsilon - 0.0005^2\right)^2}{2 \cdot 0.0002^2}}$$

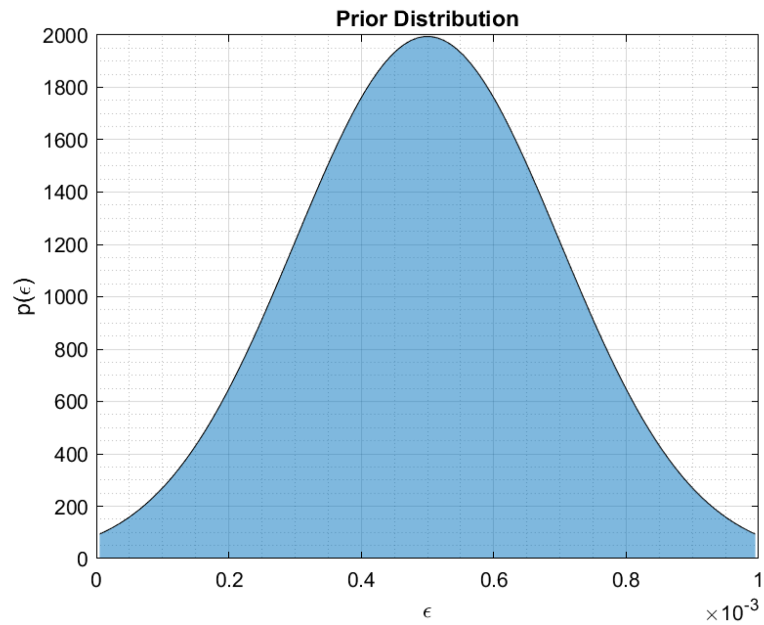

Figure 5: Prior Distribution

To compute the new posterior distribution, consider that the posterior distribution is proportional to the product of the prior distribution and the likelihood. Since we are considering the proportionality, we can drop the scale factor term on the likelihood function.

$$q\left(\boldsymbol{m}|\boldsymbol{d}\right) \propto p\left(\boldsymbol{m}\right) f\left(\boldsymbol{d}|\boldsymbol{m}\right)$$

$$\propto e^{-\frac{(\epsilon-0.0005)^2}{2\cdot 0.0002^2}} e^{\frac{-\chi^2}{2}}$$

The grid of likelihood values $f\left(\boldsymbol{d}|\boldsymbol{m}\right)$ were each multiplied by the probability density function output related to the value of ϵ associated with the grid. This resulted in a new joint probability density function which is shown in figure 6.

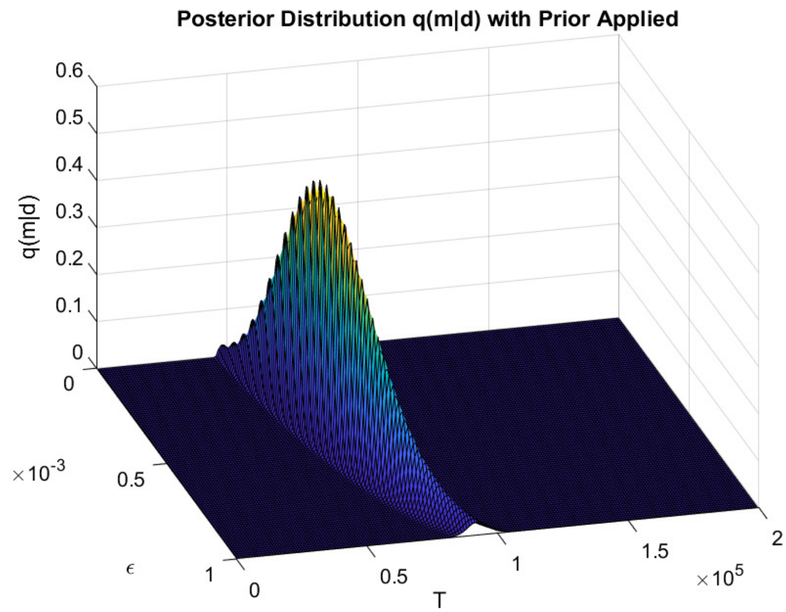

Figure 6: Posterior Distribution with Prior Applied

The newly computed marginal probabilities are also shown below in figure 7.

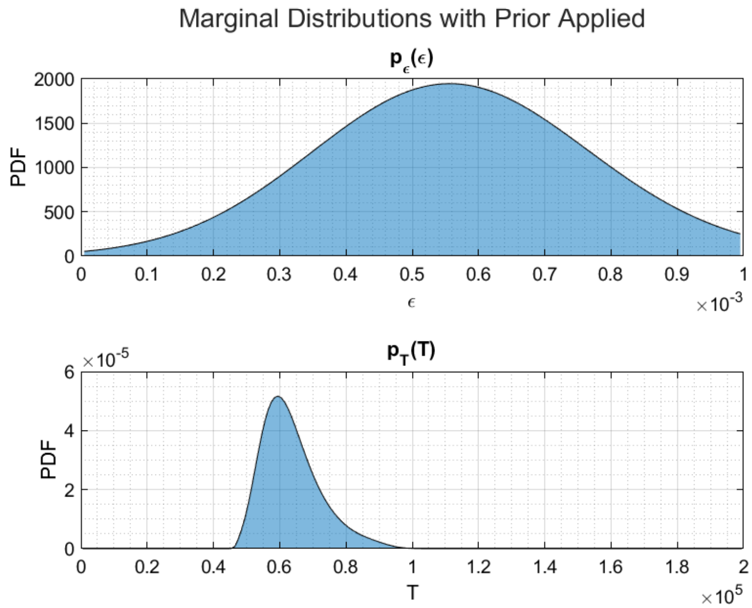

Figure 7: Marginal Distributions with Prior Applied

1.1.4 Part D

When comparing the results between using the uniformly-distributed prior versus the normally-distributed prior, it is clear that this made a key difference for the ϵ parameter. The marginal probability $p_\epsilon(\epsilon)$ *should have* appeared flat, but for some reason it showed that it was increasing. (I think there might have been a bug somewhere).

However for the normally-distributed prior case, we were able to see the mean shift in the marginal $p_\epsilon(\epsilon)$ in response to the likelihoods. It seems that the variance also grew as well. Even though the prior didn't contain any information about the T model parameter, its marginal probability density function was still reshaped into something that appears more multi-modal.

2 Problem 2

Exercise 11 in Section 11.6

2.1 Solution

2.1.1 Part A

To call the `mcmc()` function, I adjusted the function `logprior()` to `logprior1` which returns 0 if the candidate model is within our specified grid, and $-\inf$ otherwise to be consistent with a uniform distribution. I also adjusted the function `loglikelihood` to return the likelihood for a given candidate model for our problem.

I ran 10,000 samples with a step size of half of the steps used to build the grid. I tried to run more samples, but it caused my laptop to crash when I let it run over night. For an initial model, I choose to just start in the middle of the grid to see what would happen.

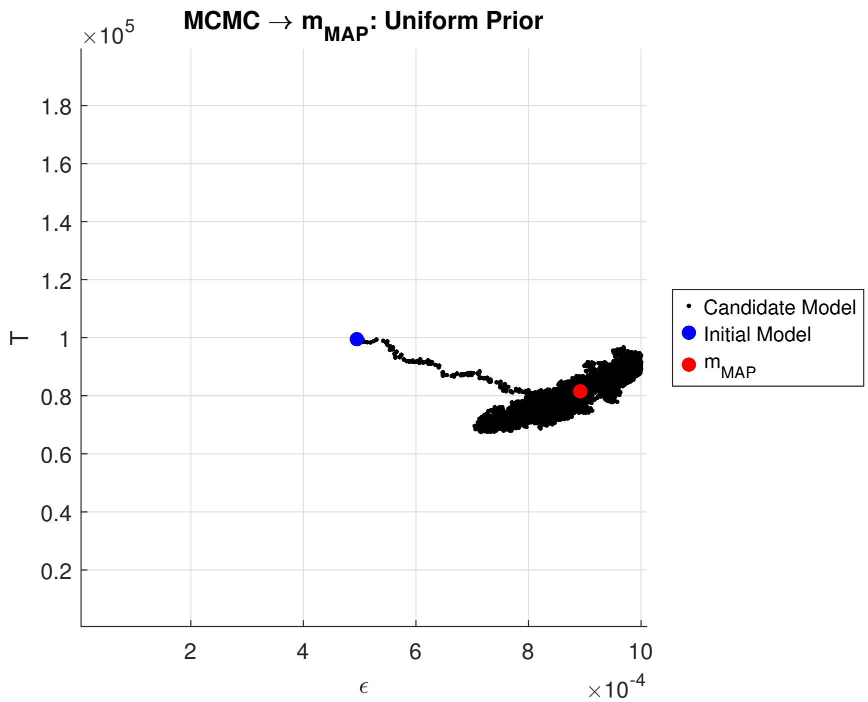

Figure 8: MCMC for a Uniform Prior

The resulting distribution from the samples appears looks similar to the posterior distribution we computed in the previous problem!

2.1.2 Part B

To approximate the marginal probabilities, I simply plotted a histogram of each row of the candidate model history from the `mcmc()` function output. In MATLAB ®, I called the `histogram()` function and used the `"pdf"` option for normalization. This resulted in the figure below.

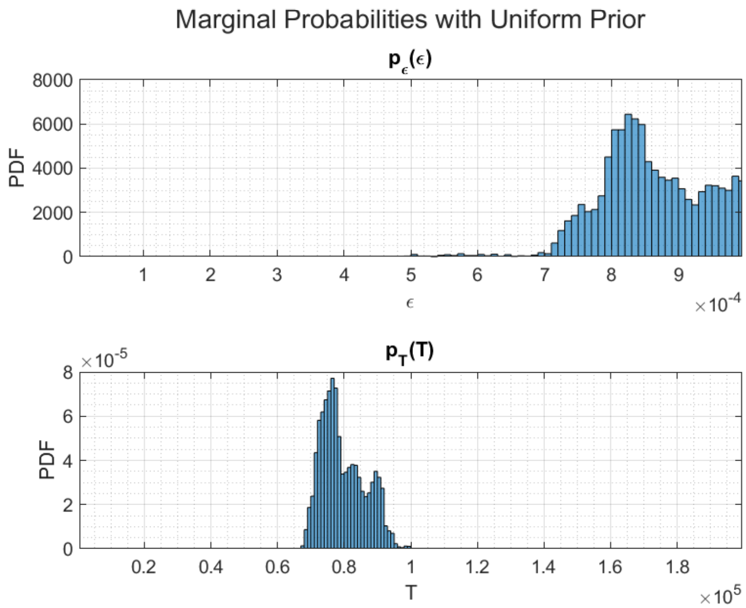

Figure 9: Marginal Probabilities with a Normal Prior

These marginal probabilities don't match the previous problem, but that might be due to a low amount of samples.

2.1.3 Part C

I repeated the same process as parts A and B, however I made a new function titled `logprior2` that computed the value of the probability density function for the candidate value of ϵ. This resulted in the figure below.

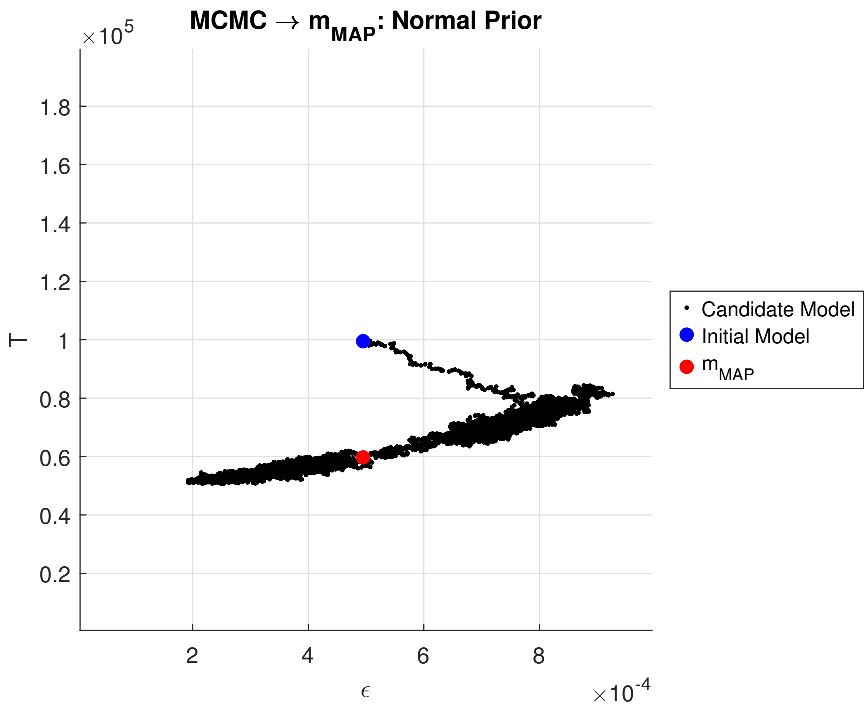

Figure 10: MCMC for a Uniform Prior

Using the same process for the marginal probabilities, I approximated the marginal probabilities below.

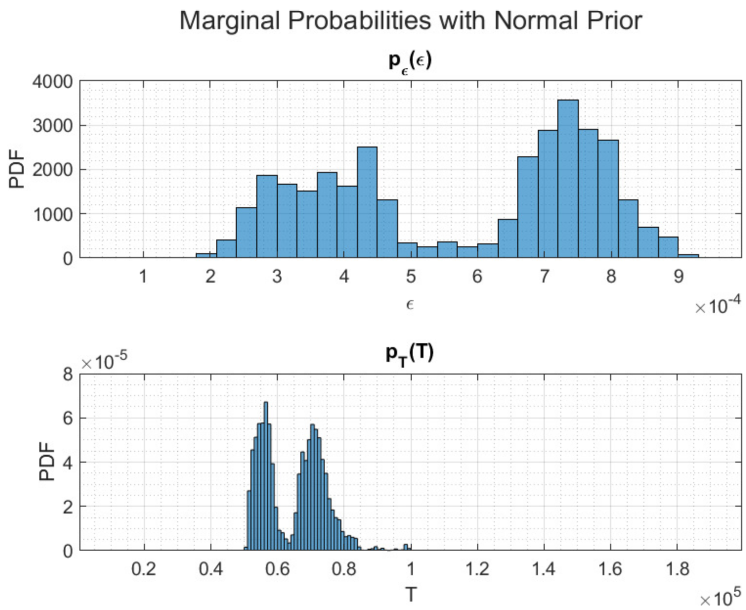

Figure 11: Marginal Probabilities with a Normal Prior

2.1.4 Part D

It took me quite a bit of time to figure out how to apply MCMC to this problem, but I am proud that I was able to figure it out. It took a lot of tuning for the number of samples and the step size to begin to see how using this MCMC method provided similar results to what we saw in the first problem.

With the uniform prior, the samples seemed to congregate outside of the original search grid that was established in the last homework assignment. The marginal probability for ϵ seems to have a similar shape as the last problem, but the histogram contains values outside of the search grid.

With the normal prior, the original scatter plot seemed to have the same distribution as the previous problem which to me appeared as a source of validation for the MCMC method being implemented properly. However, inspecting the marginal probabilities, it appears that both distributions are now bimodal, which was an unexpected result. I am curious if I can tune the step sizes or number of iterations to remove this bimodal trend if I had more time.