

CPSC 532W - Homework 3

Xiaoxuan Liang - 48131163

Public GitHub repo: <https://github.com/Xiaoxuan1121/CPSC532W/tree/main/a3>

1. Program 1: the program has been run for 50000 times, and the results are:

(a) Importance Sampling:

Importance sampling for 1.daphne took 28.951397 seconds

posterior mean of mu in 1.daphne using Importance Sampling is 7.34036922454834

posterior variance of mu in 1.daphne using Importance Sampling is 0.8557209372520447

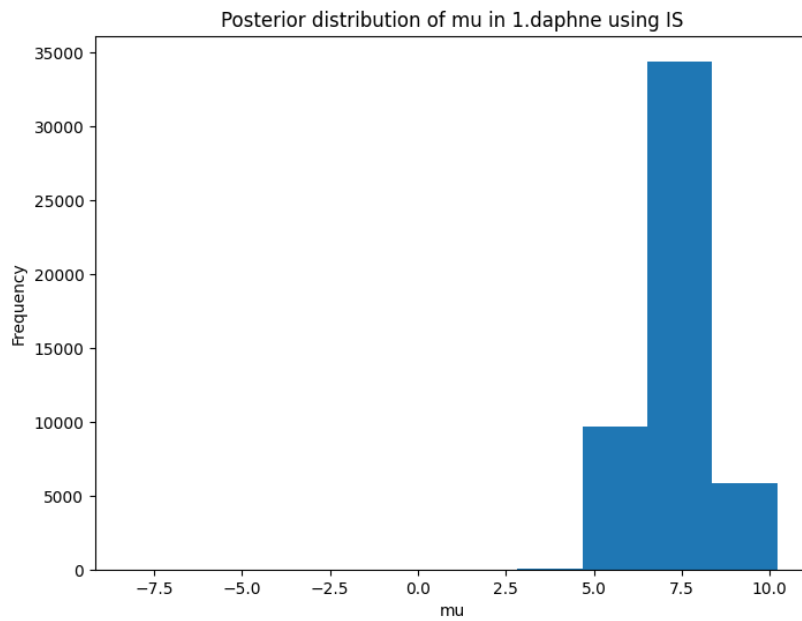


Figure 1: Posterior distribution for mu for 1.daphne using Importance Sampling

(b) MH Gibbs:

Gibbs sampling for 1.daphne took 70.520023 seconds

posterior mean of mu in 1.daphne using Gibbs is 7.228082656860352

posterior variance of mu in 1.daphne using Gibbs is 0.8902580142021179

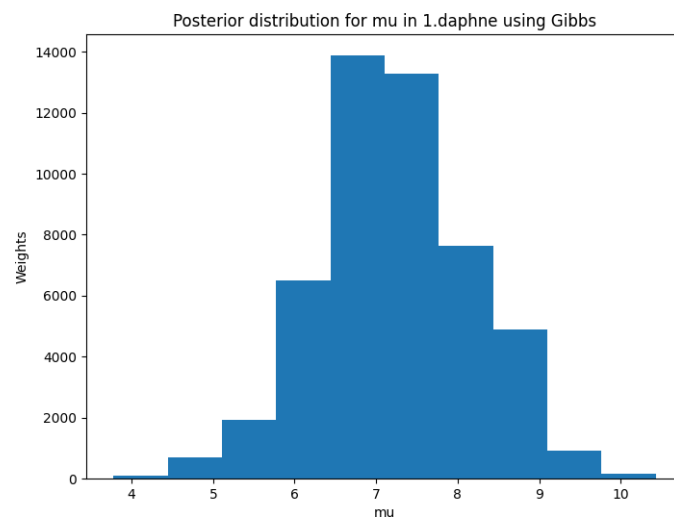


Figure 2: Posterior distribution for mu for 1.daphne using MH Gibbs Sampling

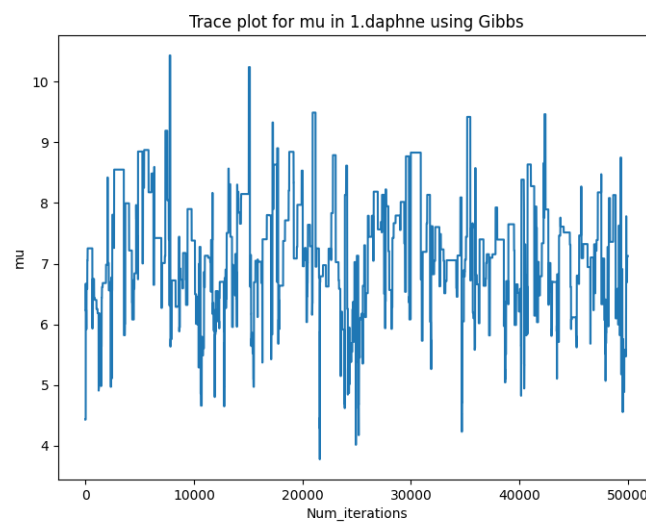


Figure 3: Trace plot for mu for 1.daphne using MH Gibbs Sampling

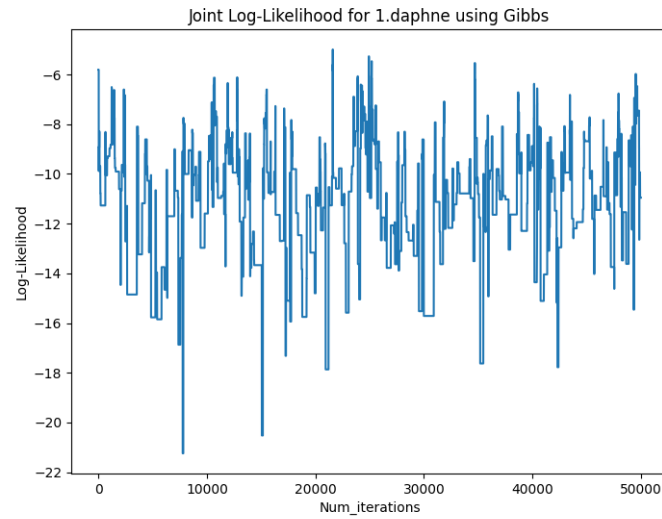


Figure 4: Joint loglikelihood plot for μ for 1.daphne using MH Gibbs Sampling

(c) HMC:

Hamiltonian monte carlo for 1.daphne took 479.338874 seconds

posterior mean of μ in 1.daphne using HMC is 7.254238605499268

posterior variance of μ in 1.daphne using HMC is 0.8068938255310059

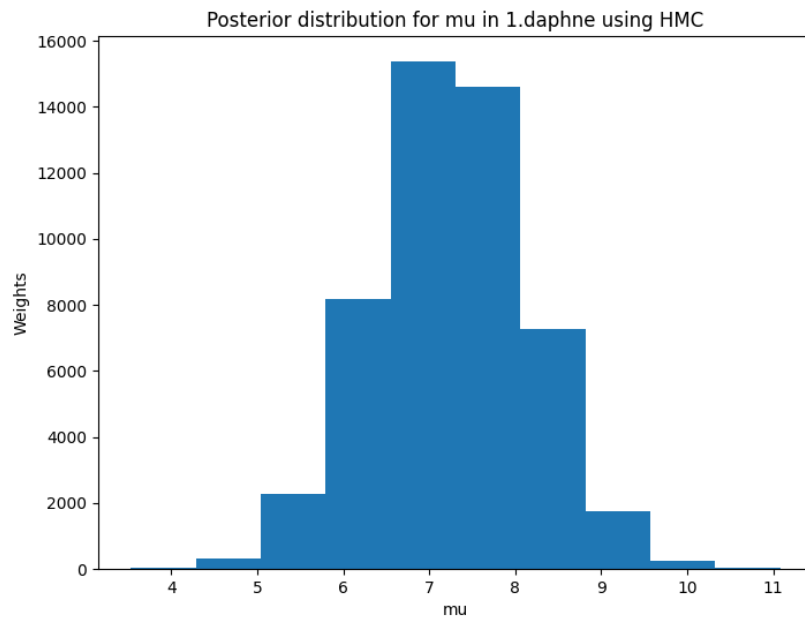


Figure 5: Posterior distribution for μ for 1.daphne using HMC

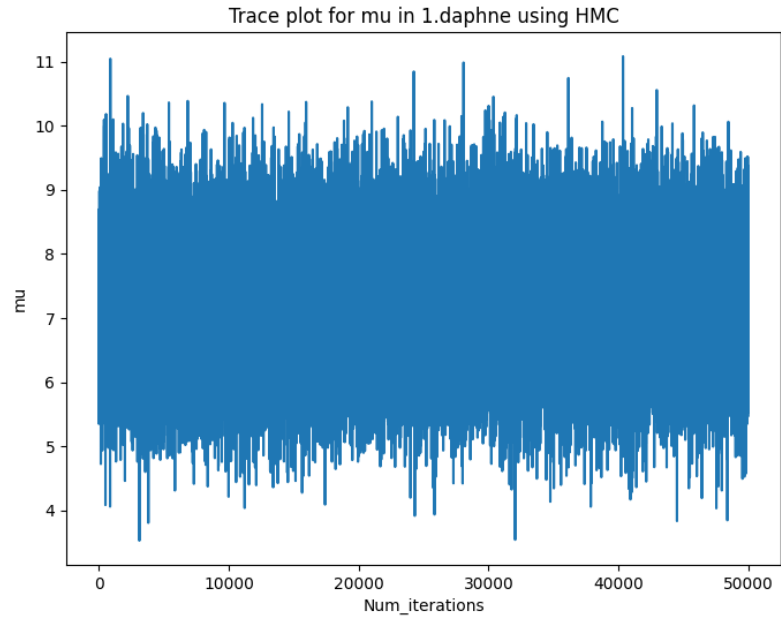


Figure 6: Trace plot for μ for 1.daphne using HMC

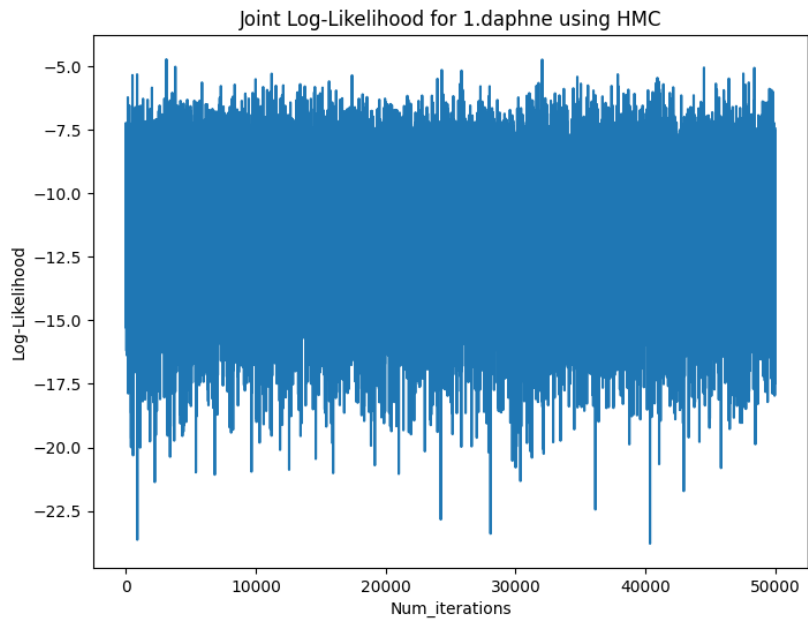


Figure 7: joint loglikelihood plot for μ for 1.daphne using HMC

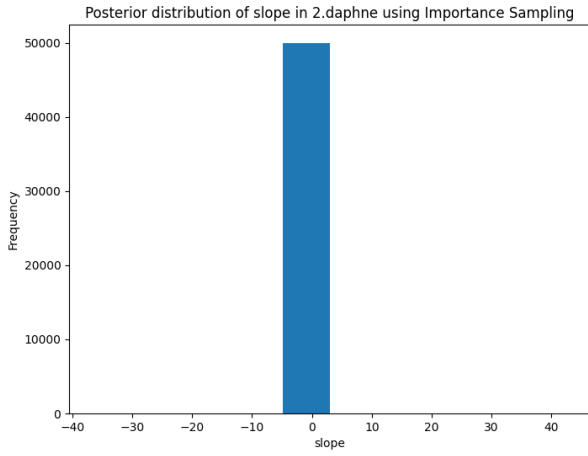
2. Program 2:

(a) Importance Sampling

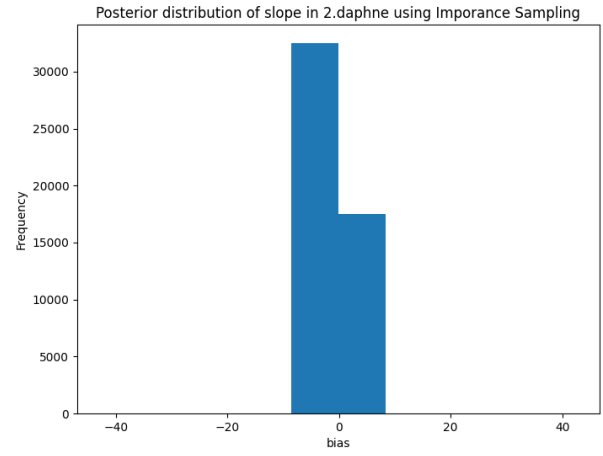
Importance sampling for 2.daphne took 77.507543 seconds

posterior means of slope and bias in 2.daphne using Importance Sampling are $\text{tensor}([2.1264, -0.3932])$

posterior covariance of slopa and bias in 2.daphne using Importance Sampling is $\begin{bmatrix} 0.05948202 & -0.20735508 \\ -0.20735508 & 0.89997893 \end{bmatrix}$



(a) Samples from the posterior for slope



(b) Samples from the posterior for bias

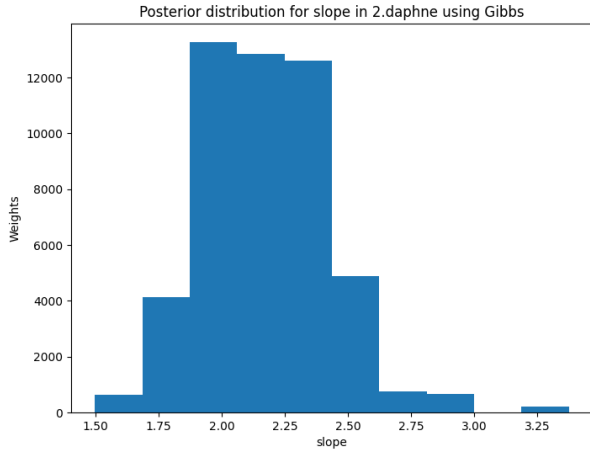
Figure 8: Posterior distribution for slope and bias for 2.daphne using Importance Sampling

(b) MH Gibbs

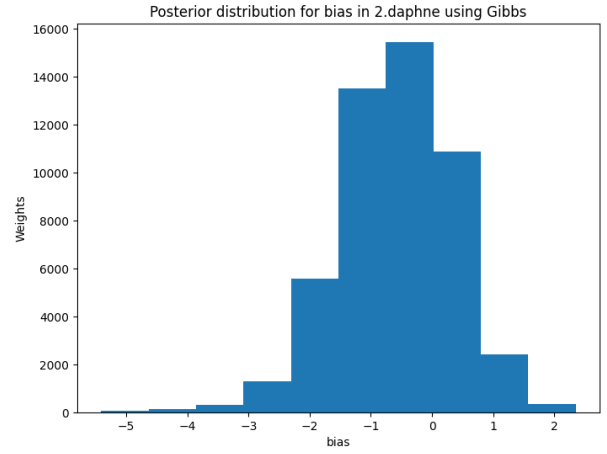
Gibbs sampling for 2.daphne took 179.560134 seconds

posterior means of slope and bias in 2.daphne using MH Gibbs Sampling is $\text{tensor}([2.1652, -0.5709])$

posterior covariance of slope and bias in 2.daphne using MH Gibbs Sampling is $\begin{bmatrix} 0.06599339 & -0.2281112 \\ -0.2281112 & 0.97806647 \end{bmatrix}$

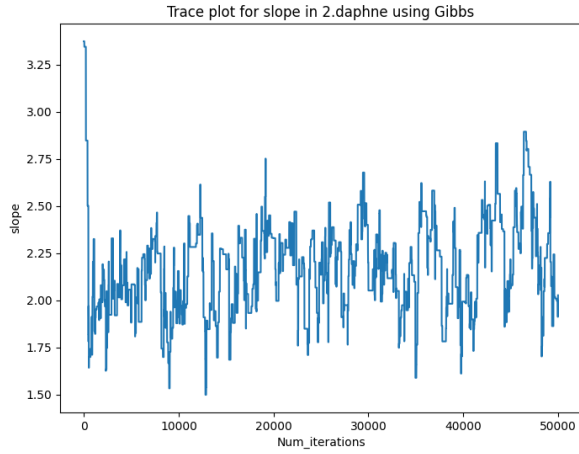


(a) Samples from the posterior for slope

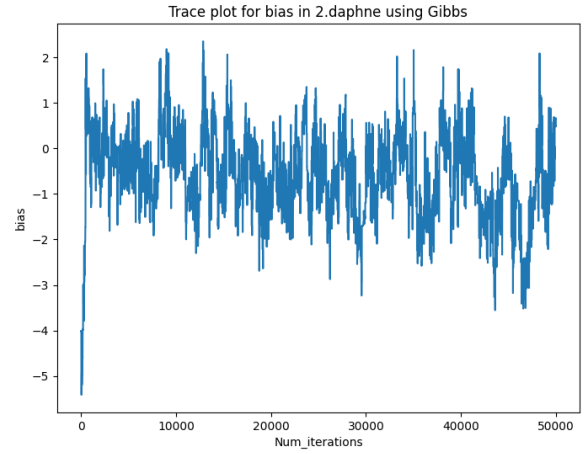


(b) Samples from the posterior for bias

Figure 9: Posterior distribution for slope and bias for 2.daphne using MH Gibbs Sampling

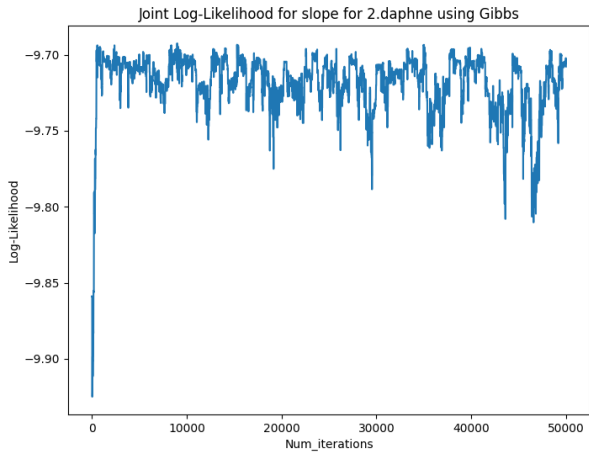


(a) Samples from the posterior for slope

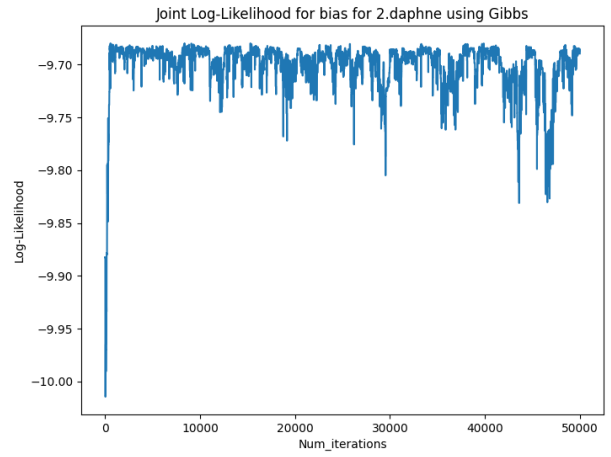


(b) Samples from the posterior for bias

Figure 10: Trace plots for slope and bias for 2.daphne using MH Gibbs Sampling



(a) Samples from the posterior for slope



(b) Samples from the posterior for bias

Figure 11: Joint Log-likelihood plots for slope and bias for 2.daphne using MH Gibbs Sampling

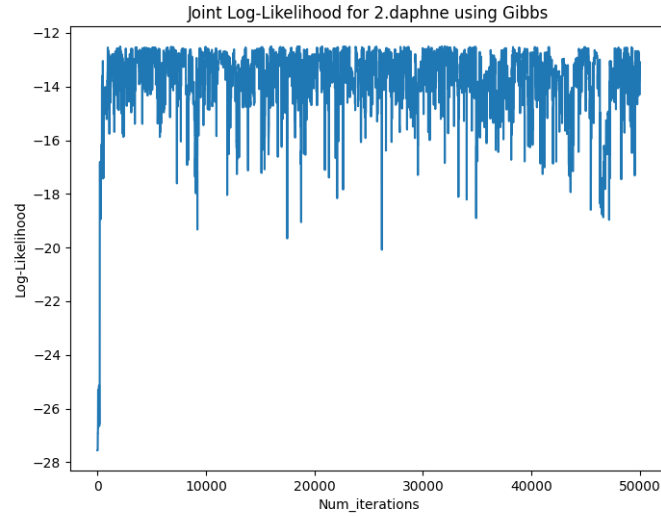


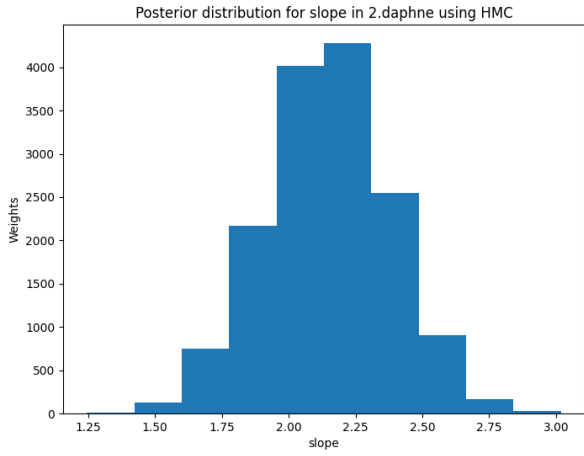
Figure 12: Joint Log-likelihood plots for 2.daphne using MH Gibbs Sampling

(c) HMC

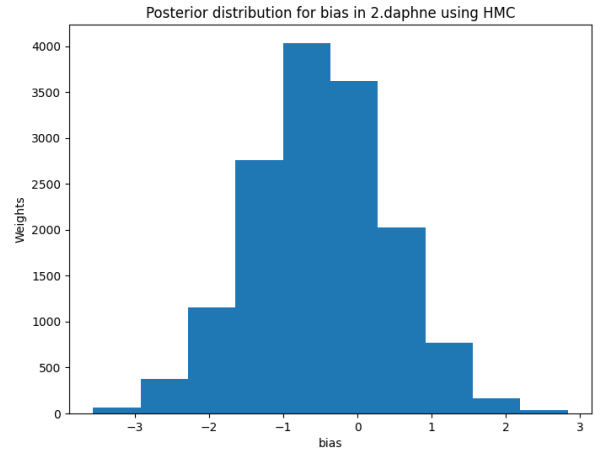
Hamiltonian monte carlo for 2.daphne took 359.792237 seconds

posterior means of slope and bias in 2.daphne using HMC is tensor([2.1515, -0.5122], grad_fn=<MeanBackward1>)

posterior covariance of slope and bias in 2.daphne using HMC is [[0.05823189 -0.18699317]
[-0.18699317 0.8151532]]

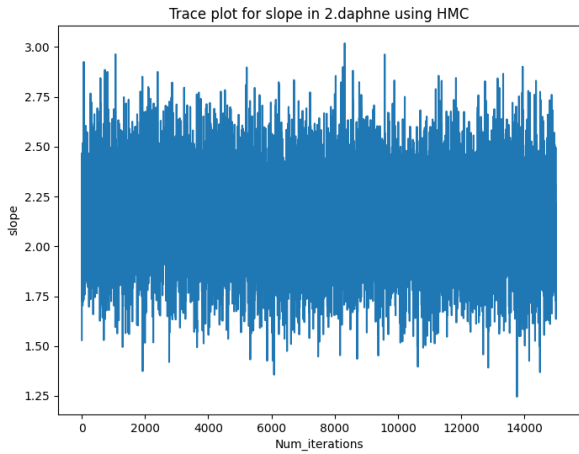


(a) Samples from the posterior for slope

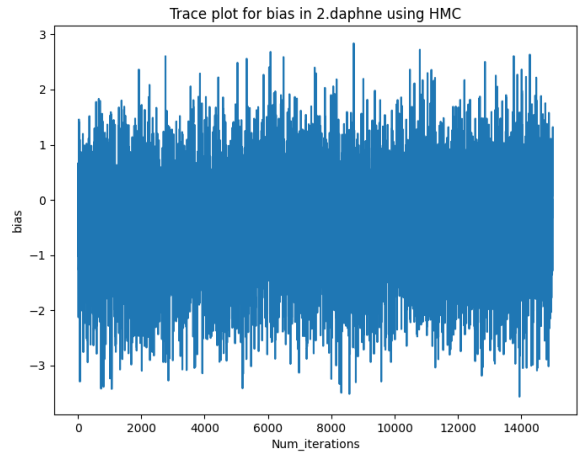


(b) Samples from the posterior for bias

Figure 13: Posterior distribution for slope and bias for 2.daphne using HMC Sampling

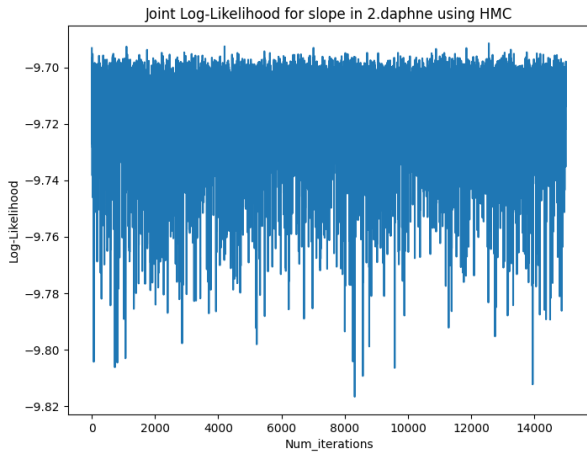


(a) Samples from the posterior for slope

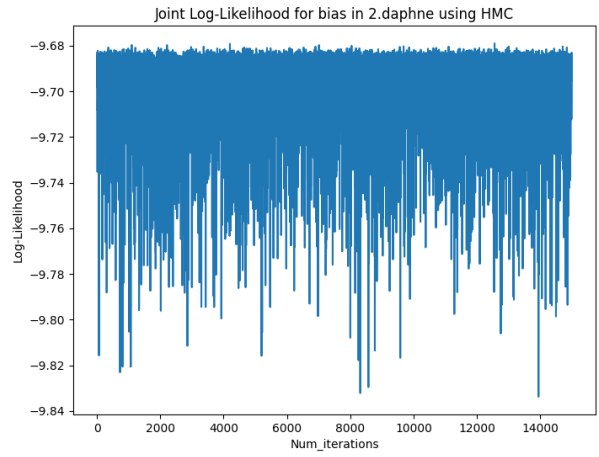


(b) Samples from the posterior for bias

Figure 14: Trace plots for slope and bias for 2.daphne using HMC Sampling



(a) Samples from the posterior for slope



(b) Samples from the posterior for bias

Figure 15: Joint Log-likelihood plots for slope and bias for 2.daphne using HMC Sampling

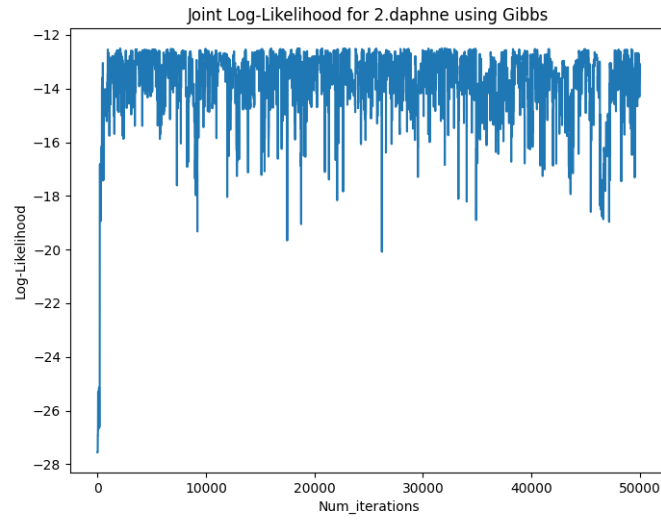


Figure 16: Joint Log-likelihood plots for 2.daphne using HMC Sampling

3. Program 3:

(a) Importance Sampling

Importance sampling for 3.daphne took 51.753358 seconds

posterior probability (mean) that the first and second datapoint are
in the same cluster using Important sampling is 0.6138077974319458

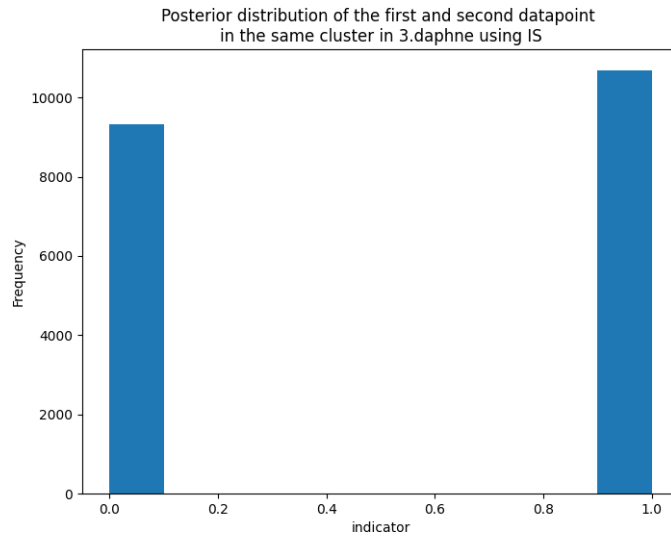


Figure 17: Posterior distribution for probability for 3.daphne using Importance Sampling

(b) MH Gibbs Sampling

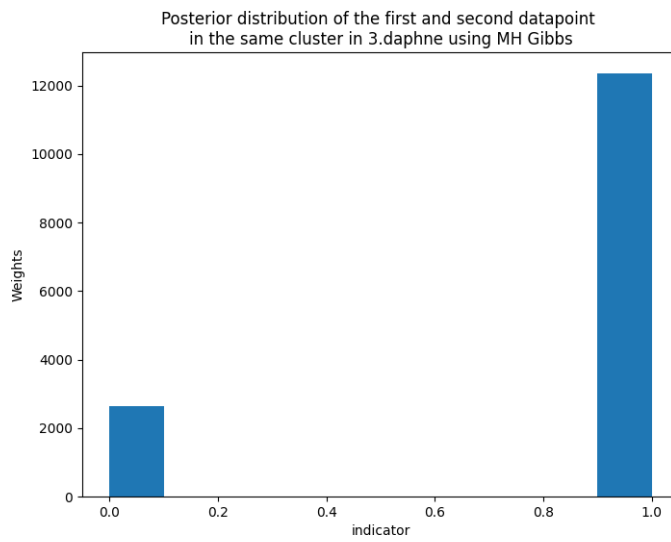


Figure 18: Posterior distribution for probability for 3.daphne using MH Gibbs Sampling

4. Program 4:

(a) Importance Sampling

Importance sampling for 4.daphne took 33.953193 seconds

posterior probability of raining in 4.daphne using Importance sampling is 0.3193276524543762

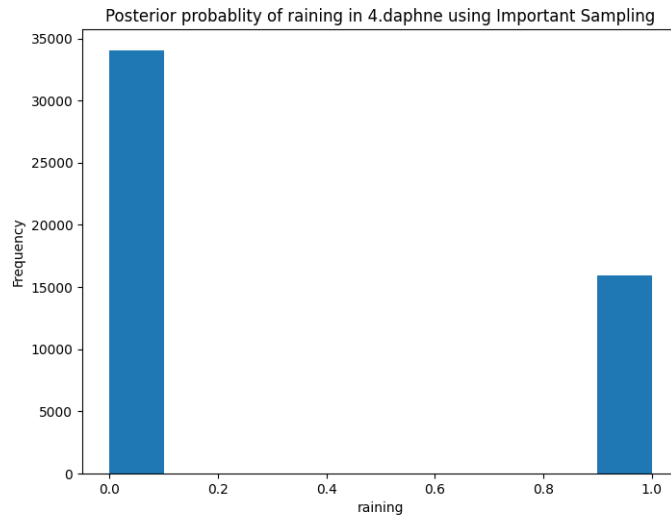


Figure 19: Posterior distribution for raining for 4.daphne using Importance Sampling

(b) MH Gibb Sampling

Gibbs sampling for 4.daphne took 212.917732 seconds

posterior probability of raining in 4.daphne using MH Gibbs is 0.3207400143146515

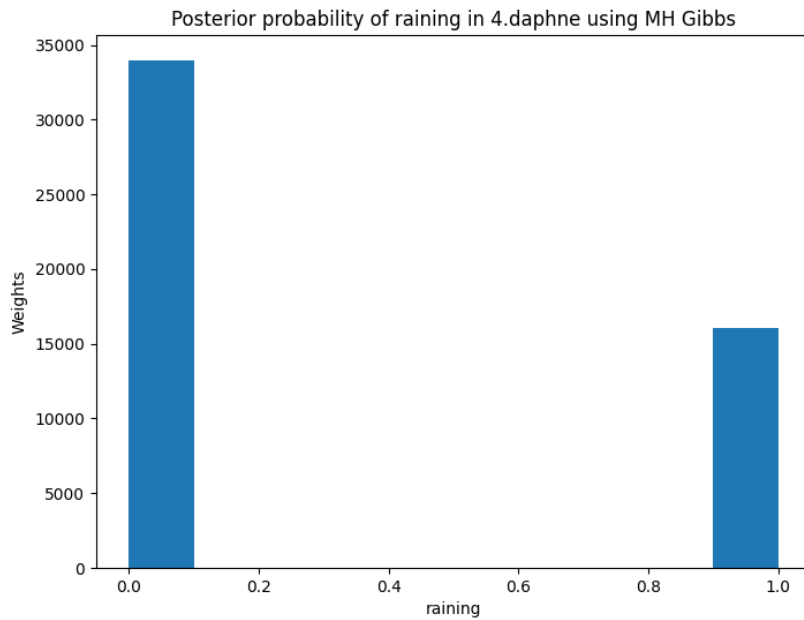


Figure 20: Posterior distribution of raining for 4.daphne using Importance Sampling

5. The caveat of directly applying Dirac distribution curing sampling process is that the likelihood goes to infinity at the center of the distribution and 0 anywhere else other than the center of the distribution, such that it is impossible to land into the region that the probability is positive. Therefore, a function is proposed to approximate Dirac distribution, which is

$$f(x) = z \exp(-(x - c)^4)$$

where c indicates the center of this function and z is a constant. It has the similar shape as what Dirac function looks like, and the highest likelihood also occurs at the center, which is several times of the likelihood corresponding to anywhere else. More important, this function is differentiable everywhere and also integrable from $-\infty$ to ∞ :

$$\int_{-\infty}^{\infty} z \exp(-(x - c)^4) dx = 2z\Gamma\left(\frac{5}{4}\right) z \Rightarrow z = \frac{1}{2\Gamma(5/4)} \text{ such that the integral is equal to 1}$$

Results analysis:

- (a) Importance Sampling seems giving a good result. Both posterior means for x and y are around 3.5 as x and y have the same priors and they are added to approximately 7.
- (b) Gibbs Sampling seems giving a not bad result. The sum of posterior means of x and y is 7 but they are not around 3.5. It makes sure that posterior means are on the line $x + y = 7$. This may due to the order that we sample x and y : we sample for x and first and update y based on the resulted value of x .
- (c) HMC gives a bad result and it's not efficient. The gradient could easily go to infinity and samples in the whole process get stuck at some numbers forever so we have to rerun it until the gradients stay at the range of finite numbers with luck. It's achievable but not reliable to calculate posterior means in this way.

(a) Importance Sampling

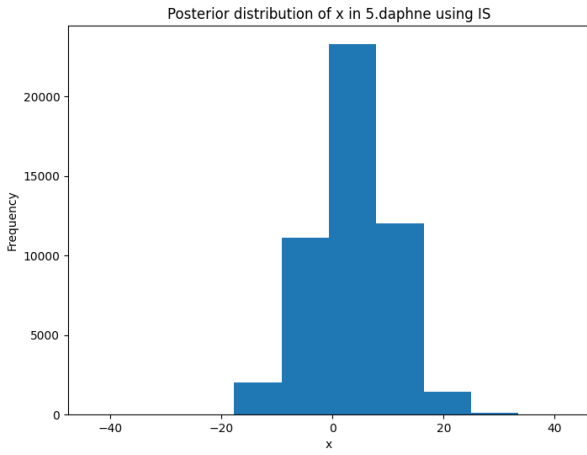
```
Importance sampling for 5.daphne took 14.455134 seconds

posterior mean of x in 5.daphne using Importance Sampling is 3.4984302520751953

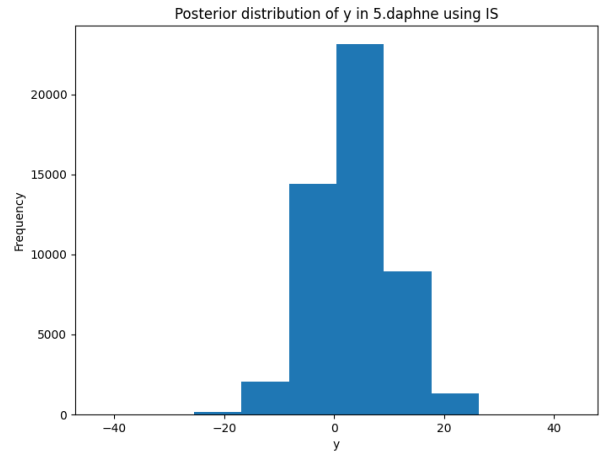
posterior mean of y in 5.daphne using Importance Sampling is 3.4777727127075195

posterior variance of x in 5.daphne using Importance Sampling is 48.890724182128906

posterior variance of y in 5.daphne using Importance Sampling is 48.88740158081055
```



(a) Samples from the prior for slope



(b) Samples from the prior for bias

Figure 21: Posterior distribution for x and y for 5.daphne using Importance Sampling

(b) MH Gibbs Sampling

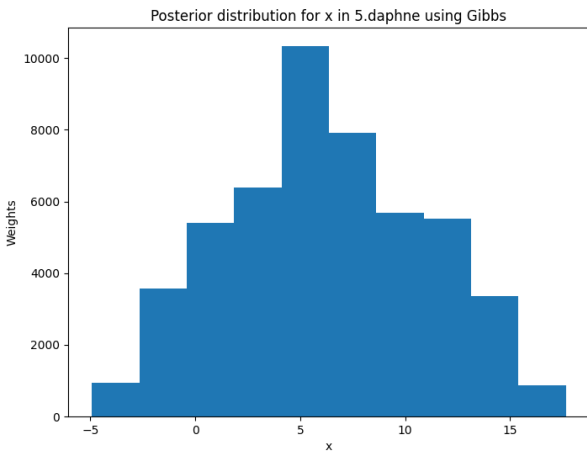
Gibbs sampling for 5.daphne took 58.316265 seconds

posterior mean of x in 5.daphne using Gibbs is 6.231391429901123

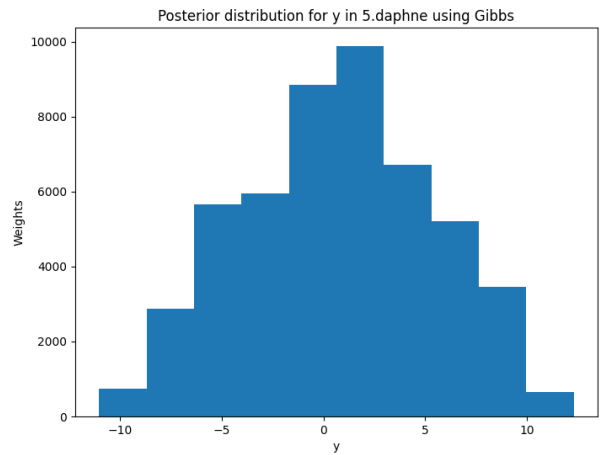
posterior mean of y in 5.daphne using Gibbs is 0.7532688975334167

posterior variance of x in 5.daphne using Gibbs is 22.28398895263672

posterior mean of x in 5.daphne using Gibbs is 0.7532688975334167

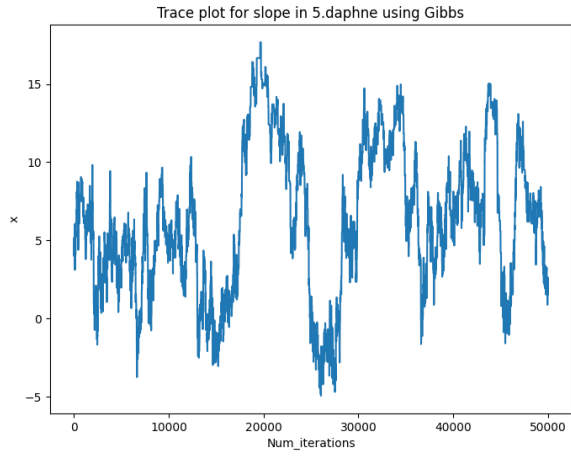


(a) Samples from the posterior for slope

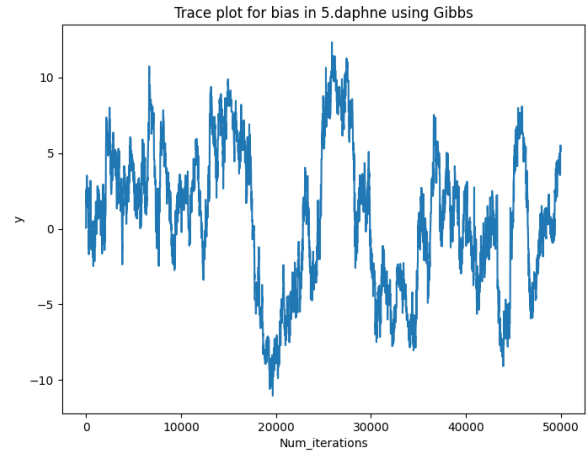


(b) Samples from the posterior for bias

Figure 22: Posterior distribution for x and y for 5.daphne using MH Gibbs Sampling

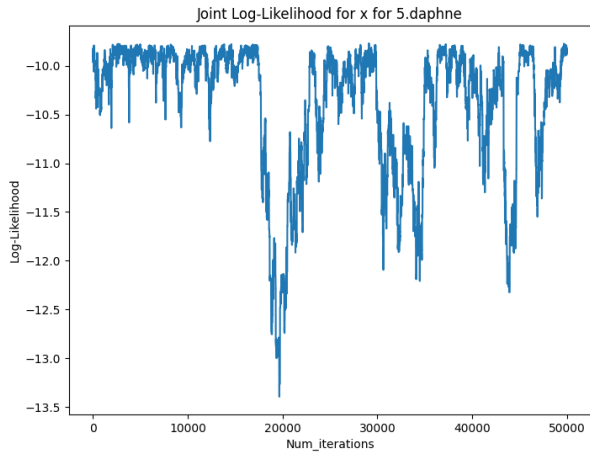


(a) Samples from the posterior for x

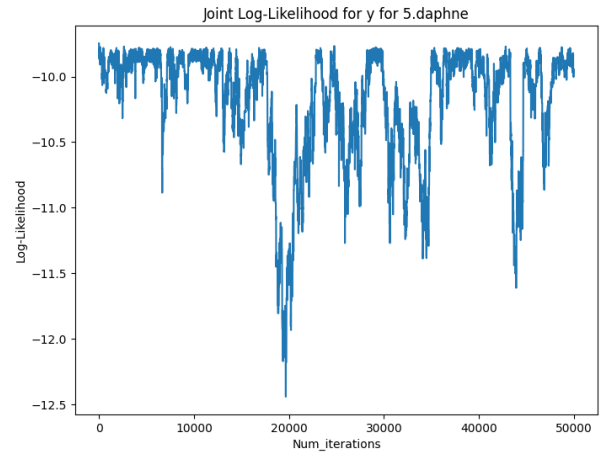


(b) Samples from the posterior for y

Figure 23: Trace plots for x and y for 5.daphne using MH Gibbs Sampling



(a) Samples from the posterior for x



(b) Samples from the posterior for y

Figure 24: Joint Log-likelihood plots for x and y for 5.daphne using MH Gibbs Sampling

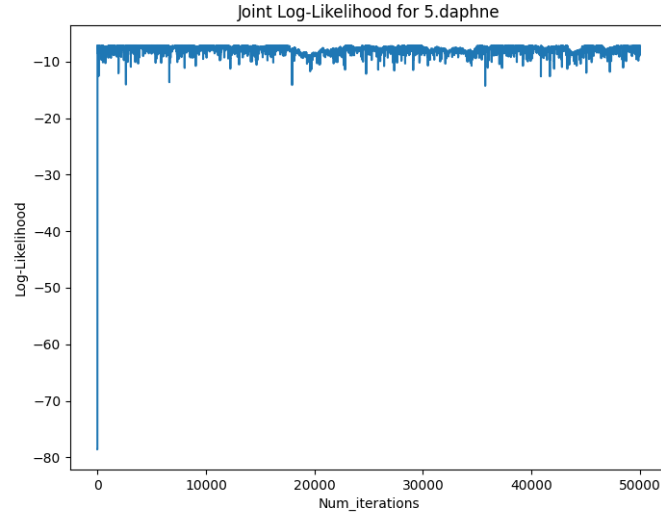


Figure 25: Joint Log-likelihood plots for 5.daphne using MH Gibbs Sampling

(c) HMC

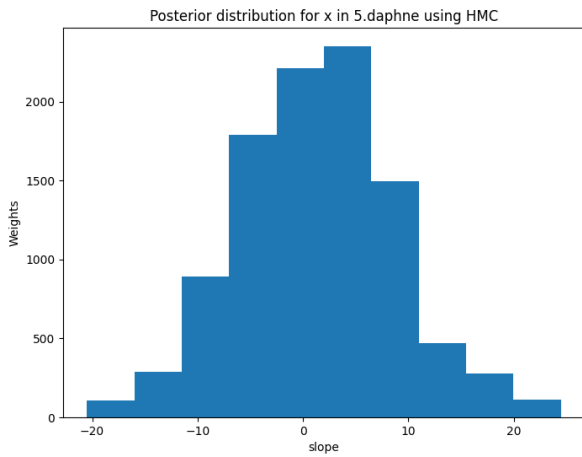
Hamiltonian monte carlo for 5.daphne took 64.211889 seconds

posterior mean of x in 5.daphne using HMC is 1.3545114994049072

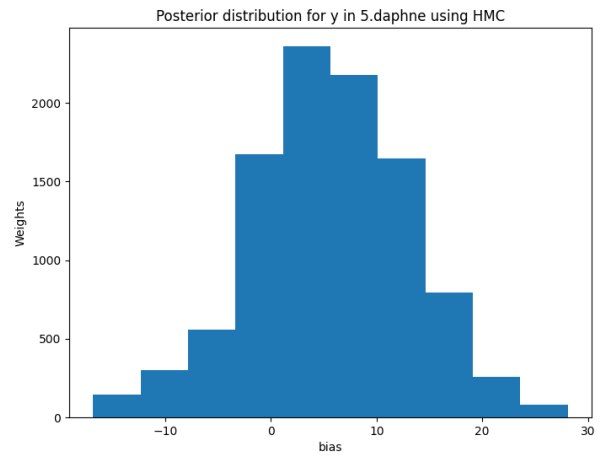
posterior mean of y in 5.daphne using HMC is 5.6317267417907715

posterior variance of x in 5.daphne using HMC is 55.87028121948242

posterior variance of y in 5.daphne using HMC is 55.87904357910156

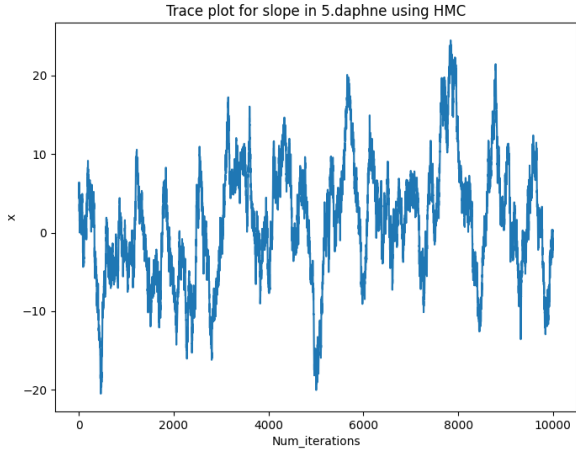


(a) Samples from the posterior for x

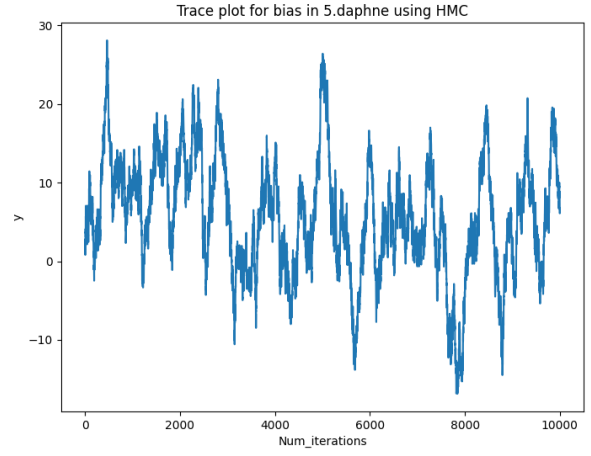


(b) Samples from the posterior for y

Figure 26: Posterior distribution for slope and bias for 5.daphne using HMC Sampling

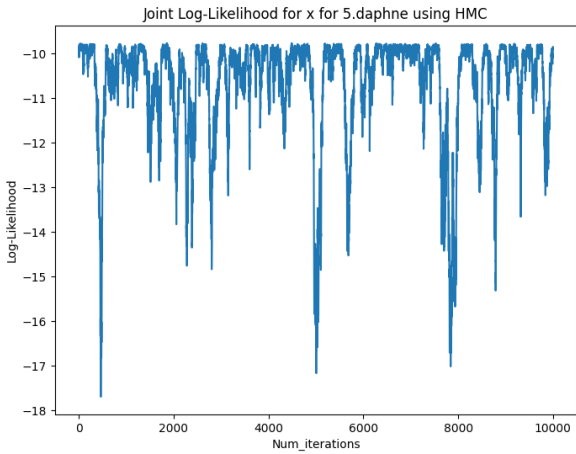


(a) Samples from the posterior for x

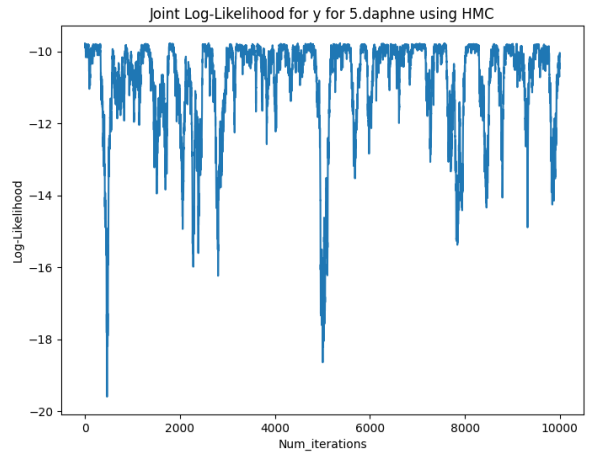


(b) Samples from the posterior for y

Figure 27: Trace plots for x and y for 5.daphne using HMC Sampling



(a) Samples from the prior for x



(b) Samples from the prior for y

Figure 28: Joint Log-likelihood plots for x and y for 5.daphne using HMC Sampling

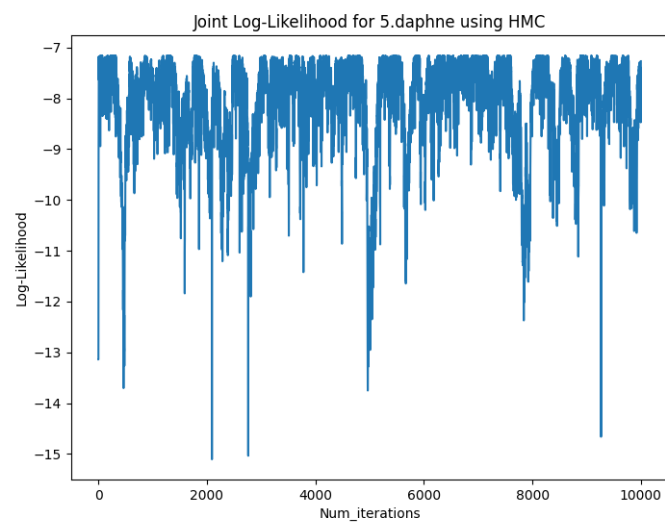


Figure 29: Joint Log-likelihood plots for 5.daphne using HMC Sampling