

Ph21 Assignment 5 - Bayesian Analysis

Yovan Badal

06/08/2018

1 Coin Toss

We use *pymc* models to set the priors and fit to a likelihood model using MCMC sampling. For the coin toss experiment, we set the true value of the probability of heads to 0.40. The results are given below:

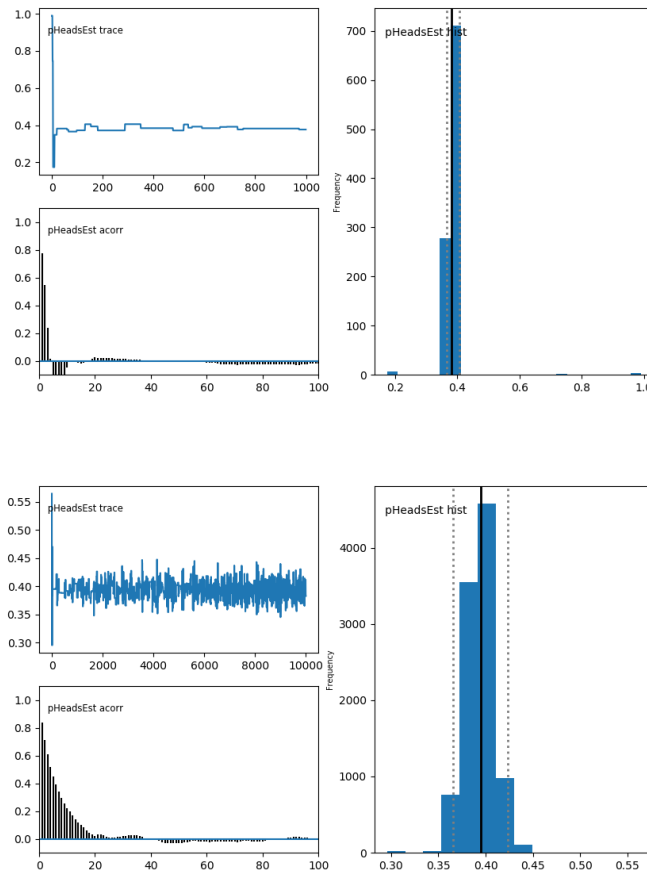


Figure 1: MCMC sampled posterior for a Uniform prior. From top to bottom: 1000 iterations, 10000 iterations.

We conduct the same simulation again, but this time we assume a Gaussian prior. The results are given below for different means and standard deviations of the priors:

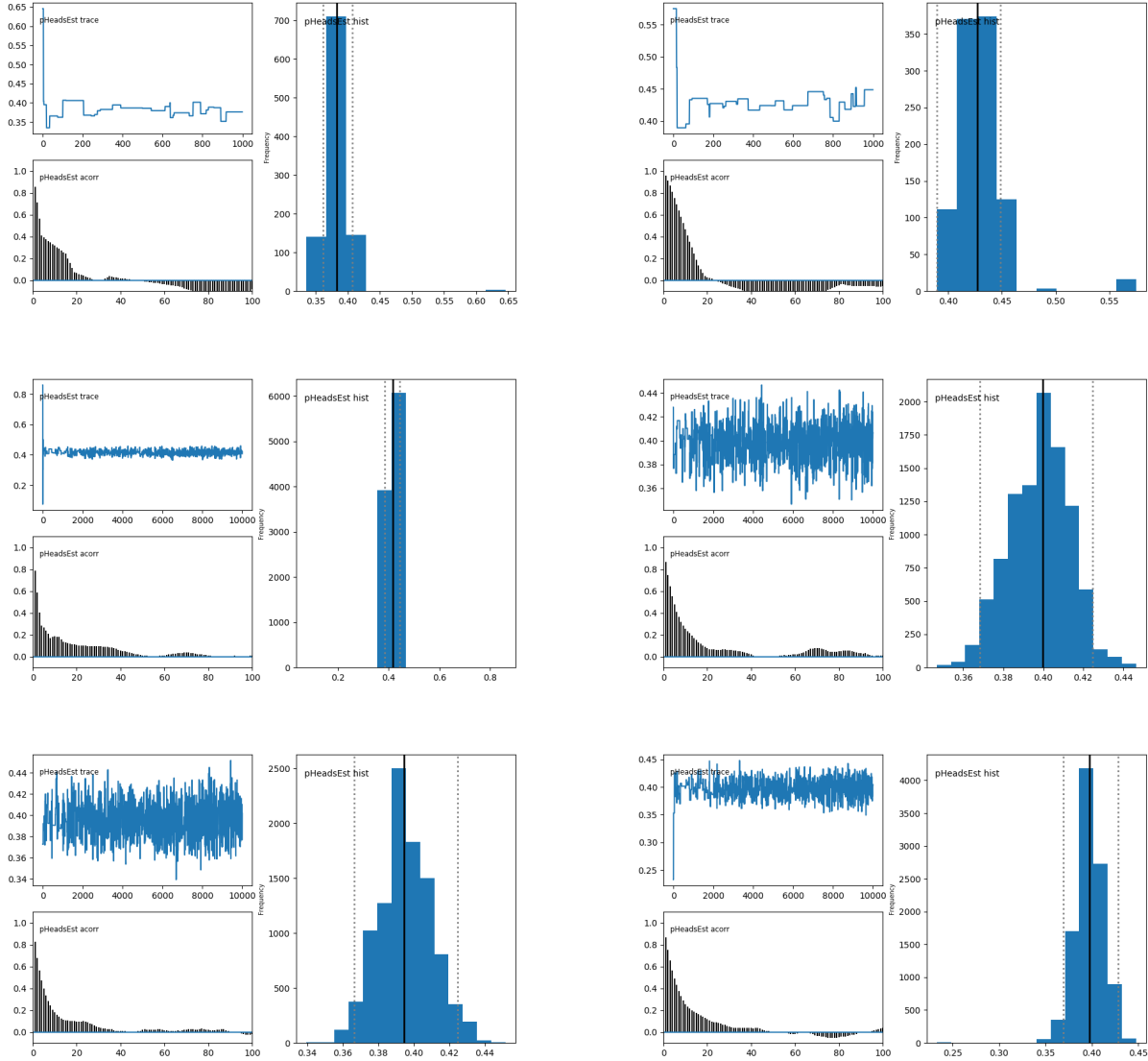


Figure 2: From right to left:

- Top row: $\{\mu = 0.5, \sigma = 0.2, 1000 \text{ iterations}\}$, $\{\mu = 0.5, \sigma = 0.05, 1000 \text{ iterations}\}$
- Middle row: $\{\mu = 0.5, \sigma = 0.2, 10000 \text{ iterations}\}$, $\{\mu = 0.5, \sigma = 0.05, 10000 \text{ iterations}\}$
- Bottom row: $\{\mu = 0.8, \sigma = 0.2, 10000 \text{ iterations}\}$, $\{\mu = 0.8, \sigma = 0.05, 10000 \text{ iterations}\}$

Observing the trace indicates that the MC distribution converges very quickly to the posterior (within 20 iterations), and as expected the mean of the sampled posterior converges much more accurately for Gaussian prior with μ within small σ of the true value of the parameter, but with large number of iterations the sampled posterior and estimated parameter are quite good nonetheless.

2 The lighthouse problem

2.1 Posterior for unknown α and β for the lighthouse as only source

We again use MCMC sampling and histogram the samples over 2 dimensions to obtain a 2d colormap of the posterior for the lighthouse problem. We set the true values $\alpha = 1.0$, $\beta = 1.5$ for the rest of our analysis, and use 50000 iterations with thinning of 5 (we drop every 5 samples to prevent auto-correlation issues) and a burn of 100 (we observe on the trace that the MC distribution converges within 100 iterations, so we discard the first 100 samples).

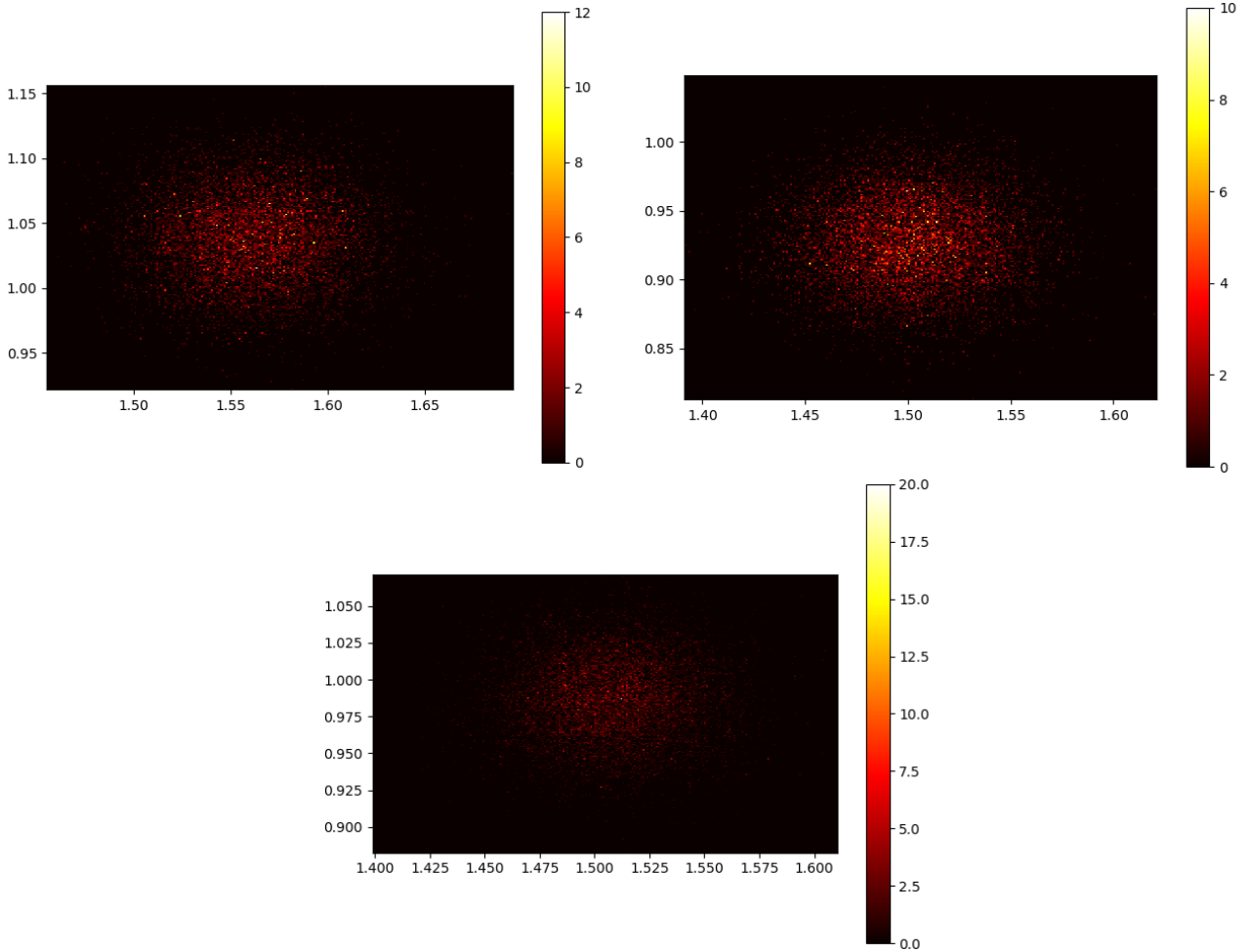


Figure 3: Colormaps with β on x-axis, α on y-axis. From right to left:

- Top right: Uniform prior
- Top left: Wide Gaussian priors
- Bottom: Narrow Gaussian priors

Again, we have much better convergence for the Gaussian prior with μ within small σ of the true values for our parameters, as expected. The centroid of the sampled posteriors however all give reasonably accurate estimates given enough iterations, although the narrow Gaussian priors generate a posterior peaking much more sharply near the true values of the parameters.

2.2 Lighthouse problem, with interloper

To the previous problem, we add an interloper (a secondary source) and run the same MCMC sampling process over the generated data. The results are shown below.

In all cases, we observe relatively poor convergence, regardless of the prior. Furthermore, the parameter estimates given by the centroid are significantly off, with as expected the estimates for the interloper far from the lighthouse being the least accurate. Therefore in this case, MCMC fails to detect the interloper without being provided with more information regarding the nature of the source (now a sum of Cauchy distributions instead of a single Cauchy), and a 4-dimensional fit is necessary.

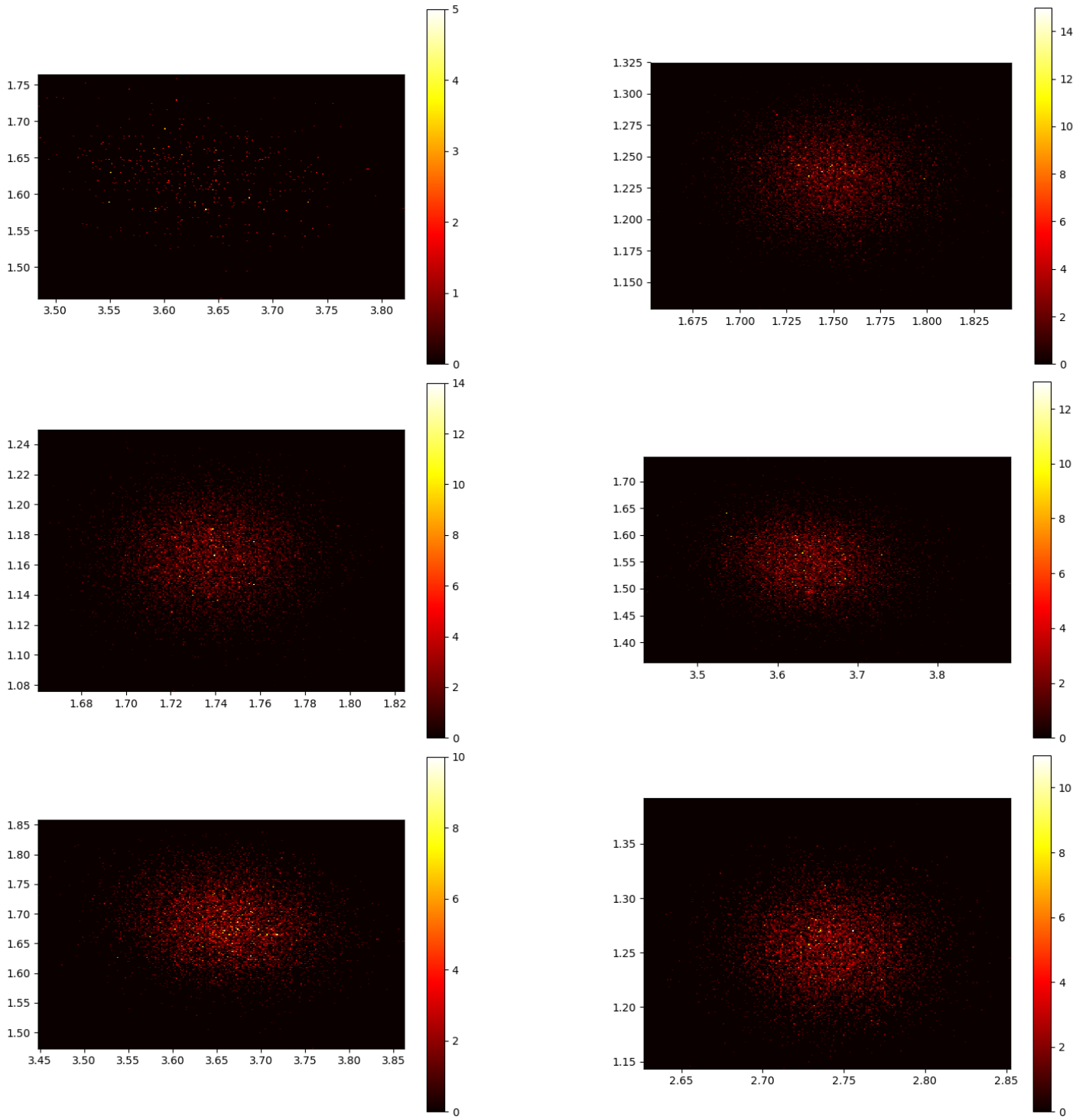


Figure 4: From right to left:

- Top row: {Uniform prior, interloper near lighthouse}, {Wide Gaussian prior, interloper near lighthouse}
- Middle row: {Narrow Gaussian prior, interloper near lighthouse}, {Uniform prior, interloper far from lighthouse}
- Bottom row: {Wide Gaussian prior, interloper far from lighthouse}, {Narrow Gaussian prior, interloper far from lighthouse}