Lab 4 Report

Chaolun Wang

October 2, 2016

1 Introduction

In this lab, the Bayes Law was used to estimate the values of parameters of an unknown distribution. With the prior mean and variance known, and a given dataset, a program was written to estimate the mean of the population successively and the confidence interval was also calculated.

2 Methodology

To solve question 1 and 2, calculate the posterior estimation of mean(denote as μ_n) is needed. Also the posterior estimation of standard deviation(denote as σ_n) is necessary for calculate the confidence interval. Thus the following math is needed:

From the given formula, we can see that :

$$p(D|\mu) \propto e^{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2}$$

also:

$$p(\mu) \propto e^{-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2}$$

multiplying these two we can have:

$$p(\mu|D) \propto p(D|\mu)P(\mu) \propto e^{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2}$$

The function above will be the posterior distribution of μ. This distribution is also a Gaussian distribution:

$$\mu|D \sim N(\mu_n, \sigma_n^2)$$

The mean μ_n will be the value of μ which maximize the value of $p(\mu|D)$, for continence I took the logarithm of $p(\mu|D)$ as the function to be maximized, I can get:

$$\frac{\partial log(p(\mu|D))}{\partial \mu} = 0$$

Which gives:

$$\frac{n}{\sigma^2}(\bar{x}-\mu) = \frac{1}{\sigma_0^2}(\mu-\mu_0)$$

In this case the μ will maximize the PDF, and it will be the mean μ_n:

$$\mu_n = \frac{\bar{x}n/\sigma^2 + \mu_0/\sigma_0^2}{1/\sigma_0^2 + n/\sigma^2}$$

The σ_n can be calculated by taking the second order partial derivative of PDF on μ. I got

$$1/\sigma_n^2 = \frac{\partial^2 log(p(\mu|D))}{\partial \mu^2} = 1/\sigma_0^2 + n/\sigma^2$$

So I have:

$$\sigma_n = \sqrt{\frac{1}{1/\sigma_0^2 + n/\sigma^2}}$$

The σ_n can then be used to calculate the confidence interval which is:

$$P(\mu_n - 0.05 < \mu < \mu_n + 0.05) = \int_{\mu_n - 0.05}^{\mu_n + 0.05} \frac{1}{\sqrt{2\pi}\sigma_n} \exp(-\frac{(\mu - \mu_n)^2}{2\sigma_n^2}) d\mu$$

Since the distribution of μ_n is also Gaussian, $\mu_n \pm Z_{\alpha/2} \times \sigma_n$ will be the confidence interval given in the formula above. The confidence coefficient $Z_{\alpha/2}$ is 1.96 if the confidence level is 95%(from online table).

Using the formulas above, I calculate the posterior value of mean for different size of n using the data set provided. also the sample number to reach confidence level is calculated in the scenario of question 1 and question 2. The data of estimated mean was generated using c++ and then plotted using in semi-log plot.

3 Result

Simulation with initial condition in question 1($\mu_0 = 0.1$, $\sigma_0 = 0.1$) and question 2($\mu_0 = 0.1$, $\sigma_0 = 1$) gives the result in Figure 1, which is plotted using the data of successive estimated mean value. We can see that both of

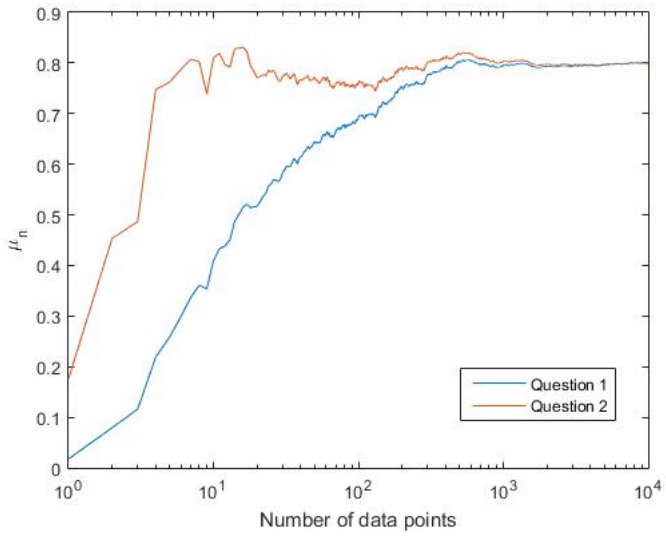

Figure 1: The successively estimated mean value as the number of data points increasing. Blue lines and red lines indicate the initial condition in question 1 and question 2 separately.

the estimations were converged to about 0.8, which is the final estimated mean given the whole data set.

For initial condition in question 1, the sample needed is 143 before the 95% confidence interval is reached for μ is ± 0.05. For initial condition in question 2, the sample needed is 153 before the 95% confidence interval reached.

4 Discussion

The differences in the priors between question 1 and question 2 is that the priors of question 2 have a larger variance. For practical implication: If the priors in question 2 is an initial guess, it can be considered that the guesser is not certain about if the mean is 0 compared with the situation in question 1, so that the mean can distribute around 0 at a larger region. By changing the variance to be larger, the guesser will be less confidence about the initial value of mean, thinking that the mean can be a little far away around where it more likely to be(which is 0).

In this case, the guesser's opinion can be easily changed and not stable(compared with question 1 situation) when the new facts are given. This situation is exactly the same as in the simulation: we can see that from the result in Figure 1, the case with larger initial variance increase faster when the data is given, and have a stronger oscillation compared with the situation in question 1.