Homework 2 Submission

Timotius Andrean Patrick Lagaunne

February, 4 2016

1. Example Run Code

```
['0.51']
Frequentist/Maximum Likelihood Probability of Heads:0.505163464892
Bayesian/MAP Probability of Heads:0.502581732446
Hit enter to continue...

['0.51', '0.34']
Frequentist/Maximum Likelihood Probability of Heads:0.421447597262
Bayesian/MAP Probability of Heads:0.447631731508
Hit enter to continue...

['0.51', '0.34', '0.31']
Frequentist/Maximum Likelihood Probability of Heads:0.385826676738
Bayesian/MAP Probability of Heads:0.414370007553
Hit enter to continue...

['0.51', '0.34', '0.31', '0.51']
Frequentist/Maximum Likelihood Probability of Heads:0.417854830856
Bayesian/MAP Probability of Heads:0.434283864685
Hit enter to continue...

['0.51', '0.34', '0.31', '0.51', '0.20']
Frequentist/Maximum Likelihood Probability of Heads:0.373922076053
Bayesian/MAP Probability of Heads:0.394935063377
Hit enter to continue...

['0.51', '0.34', '0.31', '0.51', '0.20', '0.63']
Frequentist/Maximum Likelihood Probability of Heads:0.416213485298
Bayesian/MAP Probability of Heads:0.428182987398
Hit enter to continue...

['0.51', '0.34', '0.31', '0.51', '0.20', '0.63', '0.80']
Frequentist/Maximum Likelihood Probability of Heads:0.471410882599
Bayesian/MAP Probability of Heads:0.474984522274
Hit enter to continue...

['0.51', '0.34', '0.31', '0.51', '0.20', '0.63', '0.80', '0.55']
Frequentist/Maximum Likelihood Probability of Heads:0.481135231286
Bayesian/MAP Probability of Heads:0.483231316699
Hit enter to continue...

['0.51', '0.34', '0.31', '0.51', '0.20', '0.63', '0.80', '0.55', '0.23']
Frequentist/Maximum Likelihood Probability of Heads:0.453499808142
Bayesian/MAP Probability of Heads:0.458149827328
Hit enter to continue...

['0.51', '0.34', '0.31', '0.51', '0.20', '0.63', '0.80', '0.55', '0.23', '0.79']
Frequentist/Maximum Likelihood Probability of Heads:0.487379551738
Bayesian/MAP Probability of Heads:0.488526865217
Hit enter to continue...
```

In the homework code, I use 0.5 as the true mean and the prior mean. The picture of the running of the homework2.py code that shown previously suggest that ML and MAP result differ much more in the beginning only, as the number of data increase. The MAP result will be pretty much similar with ML result because the likelihood will overcome prior in the long run. Theoretically, MAP is used when we know the prior distribution. For example, coins logically will have probability of 0.5 which is why the information and understanding that coin will logically have probability of 0.5 better to be accommodate in the model.

2. Derivation of the MAP solution

$$P(\mu|D) = \frac{P(D|\mu) \times P(\mu)}{P(D)} \quad Bayes\ Rule$$

$$P(\mu|D) \propto P(D|\mu) \times P(\mu)$$

$$P(\mu|D) \propto \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}\right) \times \left(\frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0{}^2}}\right)$$

$$P(\mu|D) \propto \left(e^{\sum_{i=1}^{n} -\frac{(x_i-\mu)^2}{2\sigma^2}}\right) \times \left(e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0{}^2}}\right) = \left(e^{\sum_{i=1}^{n} -\frac{(x_i-\mu)^2}{2\sigma^2} - \frac{(\mu-\mu_0)^2}{2\sigma_0{}^2}}\right)$$

$$\mathrm{L} = \ln[P(\mu|D)] \propto \left(\sum_{i=1}^{n} -\frac{(x_i-\mu)^2}{2\sigma^2}\right) + \left(-\frac{(\mu-\mu_0)^2}{2\sigma_0{}^2}\right)$$

When we maximize the L = $\ln[P(\mu|D)]$, we also maximize the $P(\mu|D)$.

$$\frac{dL}{d\mu} = \left(\sum_{i=1}^{n} \frac{(x_i-\mu)}{\sigma^2}\right) - \left(\frac{(\mu-\mu_0)}{\sigma_0{}^2}\right) = 0$$

$$0 = \left(\frac{(\sum_{i=1}^{n} x_i - n\mu)}{n\sigma^2}\right) - \left(\frac{(\mu-\mu_0)}{\sigma_0{}^2}\right) = \left(\frac{(\sum_{i=1}^{n} x_i)}{\sigma^2}\right) - \frac{n\mu}{\sigma^2} - \frac{\mu}{\sigma_0{}^2} + \left(\frac{(\mu_0)}{\sigma_0{}^2}\right)$$

$$\mu\left(\frac{n}{\sigma^2}+\frac{1}{{\sigma_0}^2}\right)= \left(\frac{(\sum_{i=1}^{n} x_i)}{\sigma^2}\right)+\left(\frac{(\mu_0)}{{\sigma_0}^2}\right)$$

$$\mu\left(\frac{n}{\sigma^2}+\frac{1}{{\sigma_0}^2}\right)= \left(\frac{(\sum_{i=1}^{n} x_i)}{\sigma^2}\right)+\left(\frac{(\mu_0)}{{\sigma_0}^2}\right)$$

Since the variance is assumed to be fix then

$$\mu(n+1)= \left(\sum_{i=1}^{n} x_i\right)+(\mu_0)$$

$$\mu=\frac{\left((\sum_{i=1}^{n} x_i)\right)+(\mu_0)}{n+1}$$

Given m= the number of toss that give head then

$$\mu=\frac{(m)+(\mu_0)}{n+1}$$

Is the mean solution of the MAP