# HW: Bayesian Parameter Estimation

Wen-Han Hu (whu24)

Assumption:

- Univariate Case: The data $\boldsymbol{X}$={$x_t$}, t=1,...,n is the univariate data, with the i.i.d. samples.
- Gaussian (Normal) Distribution: The sample is drawn from the Gaussian (Normal) distribution, $\boldsymbol{p(\mu) \sim N(\mu, \sigma^2)}$, with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\sigma^2}$.
- Parameters: Unknown mean, known variance
- Priors: The conjugate prior for $\boldsymbol{\mu}$ is Gaussian, $\boldsymbol{p(\mu) \sim N(\mu_0, {\sigma_0}^2)}$

*Question1and Question 2:*

Based on the assumption above, we can have the prior density and likelihood density as below:

$$\text{Prior: } p(\mu) \sim \frac{1}{\sqrt{2\pi{\sigma_0}^2}} exp\left(\frac{-1}{2{\sigma_0}^2}(\mu - \mu_0)^2\right)$$

$$\text{Likelihood: } p(\text{X}|\mu) \sim \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi{\sigma_0}^2}} exp(\frac{-1}{2\sigma^2}(x^i - \mu)^2)$$

$$\text{Posterior: } p(\mu|X) \sim p(\text{X}|\mu)\ p(\mu)$$

Since the variance is known, the constant can be removed, the Prior and Likelihood can be re-written as:

$$p(X|\mu) \sim exp\left(\frac{-1}{2\sigma^2}\sum_i (x_i - \mu)^2\right)$$

$$p(\mu) \sim exp\left(\frac{-1}{2{\sigma_0}^2}(\mu - \mu_0)^2\right)$$

And Posterior can be rewritten as:

$$p(\mu|X) \sim exp\left(\frac{-1}{2\sigma^2}\sum\nolimits_i (x_i - \mu)^2\right)\ exp\left(\frac{-1}{2{\sigma_0}^2}(\mu - \mu_0)^2\right)$$

Extend the equation we can get

$$p(\mu|X) \sim exp\left(\frac{-1}{2\sigma^2}(\mu^2 - 2\mu\mu_0 + {\mu_0}^2) + \frac{-1}{2{\sigma_0}^2}\sum\nolimits_i ({x_i}^2 - 2x_i\mu + \mu^2)\right)$$

$$p(\mu|X) \sim exp\left(\frac{-1}{2}\left(\mu^2\left(\frac{1}{{\sigma_0}^2} + \frac{n}{\sigma^2}\right) - 2\mu\left(\frac{\mu_0}{{\sigma_0}^2} + \frac{\sum_i x_i}{\sigma^2}\right) + \frac{{\mu_0}^2}{{\sigma_0}^2} + \frac{\sum_i {x_i}^2}{\sigma^2}\right)\right)$$

Now, we need to get the format as Gaussian as below:

$$p(\mu|X) \sim exp\left(\frac{-1}{2{\sigma_n}^2}(\mu^2 - 2\mu\mu_n + {\mu_n}^2)\right)$$

We can remove the constant again and comparing the terms of $\mu^2$ and $-2\mu\mu_n$, then get:

$$\frac{-1}{2\sigma_n{}^2} = \frac{-1}{2}\left(\frac{1}{\sigma_0{}^2} + \frac{n}{\sigma^2}\right)$$
$$\sigma_n{}^2 = \frac{\sigma^2\sigma_0{}^2}{\sigma^2 + n\sigma_0{}^2}$$

With the equations above, now we can get $\mu_n$:

$$\mu_n = \sigma_n{}^2\left(\frac{\mu_0}{\sigma_0{}^2} + \frac{n\bar{x}}{\sigma^2}\right) \text{ where } \bar{x} = mean\ value\ of\ X,\ \ \textstyle\sum_i x_i \text{ rewritten as } n\bar{x}$$

Thus, the Gaussian Distribution can be derived from the equations above.
$p(\mu|X) \sim exp\left(\frac{-1}{2\sigma_n{}^2}(\mu^2 - 2\mu\mu_n + \mu_n{}^2)\right)$*, where* $\mu_n = \sigma_n{}^2\left(\frac{\mu_0}{\sigma_0{}^2} + \frac{n}{\sigma^2}\right) and\, \sigma_n{}^2 = \frac{\sigma^2\sigma_0{}^2}{\sigma^2+n\sigma_0{}^2}$

*Question 3:*

From the previous equation, we can further extend $\mu_n$ :

$$\mu_n = \sigma_n{}^2\left(\frac{\mu_0}{\sigma_0{}^2} + \frac{n\bar{x}}{\sigma^2}\right) = \left(\frac{\sigma^2\mu_0 + n\bar{x}\sigma_0{}^2}{n\sigma_0{}^2 + \sigma^2}\right)$$

And $1/\sigma^2$ can be rewritten as:

$$\frac{1}{\sigma_n{}^2} = \left(\frac{1}{\sigma_0{}^2} + \frac{n}{\sigma^2}\right)$$

*Question 4:*
From the previous equations, we can identify the weights as below:

$$\mu_n = \left(\frac{n\sigma_0{}^2}{n\sigma_0{}^2 + \sigma^2}\right)\bar{x} + \left(\frac{\sigma^2}{n\sigma_0{}^2 + \sigma^2}\right)\mu_0$$

*Question 5:*
The weights are inversely proportional to the corresponding variances since $\bar{x}$ directly proportional to $\sigma_0{}^2$ and $\mu_0$ directly proportional to $\sigma^2$. The ratio of the weights is $\sigma_0{}^2/\sigma^2$

*Question 6:*
We can simply justify by adding these two weights together:

$$sum\ of\ weights = \left(\frac{n\sigma_0{}^2 + \sigma^2}{n\sigma_0{}^2 + \sigma^2}\right) = 1$$

Thus, the sum is up to 1.

*Question 7:*

The weights will always between 0 to 1. Since the sum of weights is 1, when one goes close to 1 the other one will close to 0.

*Question 8:*
Because the two weights are between 0 to 1, the value of $\mu_n$ will be

$$Min(\bar{x}, \mu_0) \leq \mu_n \leq Max(\bar{x}, \mu_0)$$

*Question 9:*
If the $\sigma^2$ is known, then we can get the equation as follow:

$$p(x^{new}|X) = \int p(x^{new}|\mu)p(\mu|X)d\mu$$

$$p(x^{new}|X) = \int N(x^{new}|\mu, \sigma^2)N(\mu|\mu_n, {\sigma_n}^2)d\mu$$

$$p(x^{new}|X) = N(x^{new}|\mu_n, {\sigma_n}^2 + \sigma^2)d\mu$$

Thus, we can perform $p(x^{new}|X) \sim N(\mu_n, {\sigma_n}^2 + \sigma^2)$

*Question 10:*

From the previous equation, we can calculate $\mu_n, {\sigma_n}^2$ $with\ n = 20$ as below:

$${\sigma_n}^2 = \frac{\sigma^2 {\sigma_0}^2}{\sigma^2 + n{\sigma_0}^2} = \frac{1.5^2 0.8^2}{1.5^2 + 20(0.8)^2} = 0.095681$$

$$\mu_n = \left(\frac{\sigma^2 \mu_0 + n\bar{x}{\sigma_0}^2}{n{\sigma_0}^2 + \sigma^2}\right) = \frac{4(1.5)^2 + 20 \times 6(1.5)^2}{1.5^2 + 20(0.8)^2} = 5.70099$$

The plot as below: