Homework 2

Ankur Garg, agarg12@ncsu.edu

Q1.

$X = \{x^1, x^2, x^3, \ldots, x^n\}$ is data of n samples which are independent and identically distributed.

Prior Distribution for μ is known to be: $p(\mu) \sim N\ (\mu_0, \sigma_0^2)$

The samples are drawn from a Gaussian Distribution: $p(x) \sim N\ (\mu, \sigma^2)$ where the mean is unknown and the variance is known.

Using these details:

The likelihood of the data samples X for a given mean μ can be calculated as following:

$$p(X|\mu) = p(x^1|\mu) * p(x^2|\mu) * \ldots * p(x^n|\mu)$$

(Because the samples are independent and identically distributed)

Therefore, $$p(X|\mu) \sim \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) \exp\left(-\frac{1}{2\sigma^2}\left(x^i - \mu\right)^2\right)$$

$$p(X|\mu) \sim \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x^i - \mu\right)^2\right)$$

Here, variance is known so the first term in the expression is a constant and can be ignored for now.

$$p(X|\mu) \sim \exp\left(-\frac{1}{2\sigma^2}\left(\Sigma_i x^2 + n\mu^2 - 2\mu n\bar{X}\right)\right)$$

where $\bar{X}$ is the mean of all the samples in X

Again, here $\Sigma_i x^2$ is a constant term in the expression, and since we are just looking at the proportionality of $p(X|\mu)$, it can be ignored.

$$p(X|\mu) \sim \exp\left(-\frac{n}{2\sigma^2}\left(\mu^2 - 2\mu\bar{X}\right)\right)$$

With some manipulation of constants, this can be re-written as:

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

This is the likelihood distribution of the data in terms of an unknown μ

The posterior distribution of the mean, $p(\mu|X)$ can be written as follows:

$$p(\mu|X) \sim p(X|\mu) * p(\mu)$$

Substituting for likelihood and prior distributions, we would get:

$$p(\mu|X) \sim \exp\left(-\frac{1}{2\sigma^2}(\bar{X}-\mu)^2\right) * \left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right)\exp\left(-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2\right)$$

This can be re-written as:

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

This can be further simplified as:

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

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

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

Here, again the term $\left(\frac{n\bar{X}^2}{\sigma^2}+\frac{\mu_0^2}{\sigma_0^2}\right)$ is a constant and can be ignored for the purposes of this problem.

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

This is the posterior distribution of the $\mu|X$

We need to get this in the following form:

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

By comparing the coefficients between the two expressions, we get:

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

and

$$-\frac{1}{2\sigma_n^2} * \mu_n = -\frac{1}{2}\left(\frac{n\bar{X}}{\sigma^2} + \frac{\mu_0}{\sigma_0^2}\right)$$

from first equation, we get:

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

(Equation 2).

and from second one, we get:

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

substituting:

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

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

(Equation 3).

So, the posterior distribution is of the form:

$p(\mu|X) \sim \exp\left(-\frac{1}{2\sigma_n^2}(\mu - \mu_n)^2\right)$, where μ_n and σ_n^2 are as above

This completes the first three questions, for these expressions for $\mu_n$and σ_n^2, we see that the posterior distribution follows a Gaussian distribution $\sim N\ (\mu_n, \sigma_n^2)$

Q2. Proved in Q1 (Equation 1 onwards till the end of Q1).

Q3. The value of μ_n and σ_n^2 is as calculated in Q1. Check Equation 2 and Equation 3.

Reference.: For questions 1,2,3, Youtube: https://www.youtube.com/user/deetoher/videos

Q4.

Looking at the expression for μ_n, we can write it in the following way:

$$\mu_n = \frac{n\sigma_0^2}{\sigma^2 + n\sigma_0^2}\bar{X} + \frac{\sigma^2}{\sigma^2 + n\sigma_0^2}\mu_0$$

This is weighted average of the prior mean μ_0 and the sample mean $\bar{X}$

Q5.
The weights are inversely proportional to their variances. Because:
The weight of the sample mean $\bar{X}$ is directly proportional to the prior variance σ_0^2 and the weight of the prior mean μ_0 is directly proportional to the sample variance σ^2

Also, the ratio of the weights:

$$\frac{Weight\ corresponding\ to\ sample\ mean}{Weight\ corresponding\ to\ prior\ mean} = \frac{\sigma_0^2}{\sigma^2}$$

So, they are inversely proportional to the corresponding variances

Q6.
The sum of the weights here is:

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

Q7.
Yes, each weight is between 0 and 1 because:

Sample weight = $\frac{n\sigma_0^2}{\sigma^2+n\sigma_0^2} = \frac{1}{\sigma^2/n\sigma_0^2+1}$, Here the numerator is 1 and the denominator is greater than 1. So, the weight is between 0 and 1.

Similarly, the prior weight = $\frac{\sigma^2}{\sigma^2+n\sigma_0^2} = \frac{1}{1+n\sigma_0^2/\sigma^2}$, Here the numerator is 1 and the denominator is greater than 1. So, the weight is between 0 and 1.

Q8.
Based on these weights, we can infer the following about the value of μ_n w.r.t to the $\bar{X}$ and μ_0

We know that the two weights are between 0 and 1 so, in extreme cases, in one of the weights is 1 and another is 0, in that case max possible value would be equal to $\max(\bar{X}, \mu_0)$ and the min value would be $\min(\bar{X}, \mu_0)$. So,
The value of μ_n must be between $\bar{X}$ and μ_0 i.e.:

$$\min(\bar{X}, \mu_0) \leq \mu_n \leq \max(\bar{X}, \mu_0)$$

Q9.
Given variance is known, for a new data point X^{new},

$$p(x^{new}|X) = \int_{-\infty}^{\infty} p(x^{new}|\mu) * p(\mu|X)d\mu$$

Substituting from previous results:

$$p(x^{new}|X) \sim \frac{1}{2\pi\sqrt{(\sigma^2\sigma_n^2)}}\int \exp\left(-\frac{1}{2\sigma^2}(x^{new}-\mu)^2\right) * \exp\left(-\frac{1}{2\sigma_n^2}(\mu-\mu_n)^2\right)d\mu$$

$$p(x^{new}|X) \sim \frac{1}{2\pi\sqrt{(\sigma^2\sigma_n^2)}}\int \exp\left(-\frac{1}{2\sigma^2}(x^{new}-\mu)^2-\frac{1}{2\sigma_n^2}(\mu-\mu_n)^2\right)d\mu$$

$$p(x^{new}|X) \sim \frac{1}{2\pi\sqrt{(\sigma^2\sigma_n^2)}}\int \exp\left(-\frac{1}{2}\left(\frac{1}{\sigma^2}(x^{new}-\mu)^2+\frac{1}{\sigma_n^2}(\mu-\mu_n)^2\right)\right)d\mu$$

$$p(x^{new}|X) \sim \frac{1}{2\pi\sqrt{(\sigma^2\sigma_n^2)}}\int \exp\left(-\frac{1}{2}\left(\frac{x^{new\,2}+\mu^2-2\mu x^{new}}{\sigma^2}+\frac{\mu^2+\mu_n^2-2\mu\mu_n}{\sigma_n^2}\right)\right)d\mu$$

$$p(x^{new}|X) \sim \frac{1}{2\pi\sqrt{(\sigma^2\sigma_n^2)}}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left(\mu^2\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_n^2}\right)-2\mu\left(\frac{x^{new}}{\sigma^2}+\frac{\mu_n}{\sigma_n^2}\right)+\left(\frac{x^{new\,2}}{\sigma^2}+\frac{\mu_n^2}{\sigma_n^2}\right)\right)\right)$$

This can be re-written as:

$$p(x^{new}|X) \sim \frac{1}{\sqrt{2\pi(\sigma^2+\sigma_n^2)}}\exp\left(-\frac{(x^{new}-\mu_n)^2}{2(\sigma^2+\sigma_n^2)}\right)\int_{-\infty}^{\infty}\frac{\sqrt{\sigma^2+\sigma_n^2}}{\sqrt{2\pi}\sigma\sigma_n}\exp\left(-\frac{(\mu-\frac{\sigma^2(x^{new})+\sigma_n^2\mu_n}{\sigma^2+\sigma_n^2})}{2\left(\frac{\sigma\sigma_n}{\sqrt{\sigma^2+\sigma_n^2}}\right)^2}\right)$$

The expression in the integral is a normal distribution some mean and variance.
So, the integral over –infinity to infinity will be 1.
So,

$$p(x^{new}|X) \sim \frac{1}{\sqrt{2\pi(\sigma^2+\sigma_n^2)}}\exp\left(-\frac{(x^{new}-\mu_n)^2}{2(\sigma^2+\sigma_n^2)}\right)$$

Comparing it with the standard form of normal expression,

$$N\left(\mu^{new},\sigma^{new\,2}\right) \sim \exp\left(-\frac{1}{2\sigma^{new\,2}}(x-\mu^{new})^2\right)$$

$$\sigma^{new\,2} = (\sigma^2+\sigma_n^2)$$

and,

$$\mu^{new} = \mu_n$$

So,

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

Reference: Wikipedia Page:
https://en.wikipedia.org/wiki/Sum_of_normally_distributed_random_variables

Q10.

We have likelihood distribution as: p(x) ~ N(6,1.5[2]). So,

Also, the prior distribution is: p(μ) ~ N(4,0.8[2]).

p(μ|X) $\sim N(\mu_n, \sigma_n^2) \sim N(6, 1.5^2) * N(4, 0.8^2)$

n = 20

We have the expressions for the μ_n and σ_n^2, calculated from the Q3. (Equation 2 and 3)

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

and

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

Plugging in the values for all the variables in these expressions, we get:

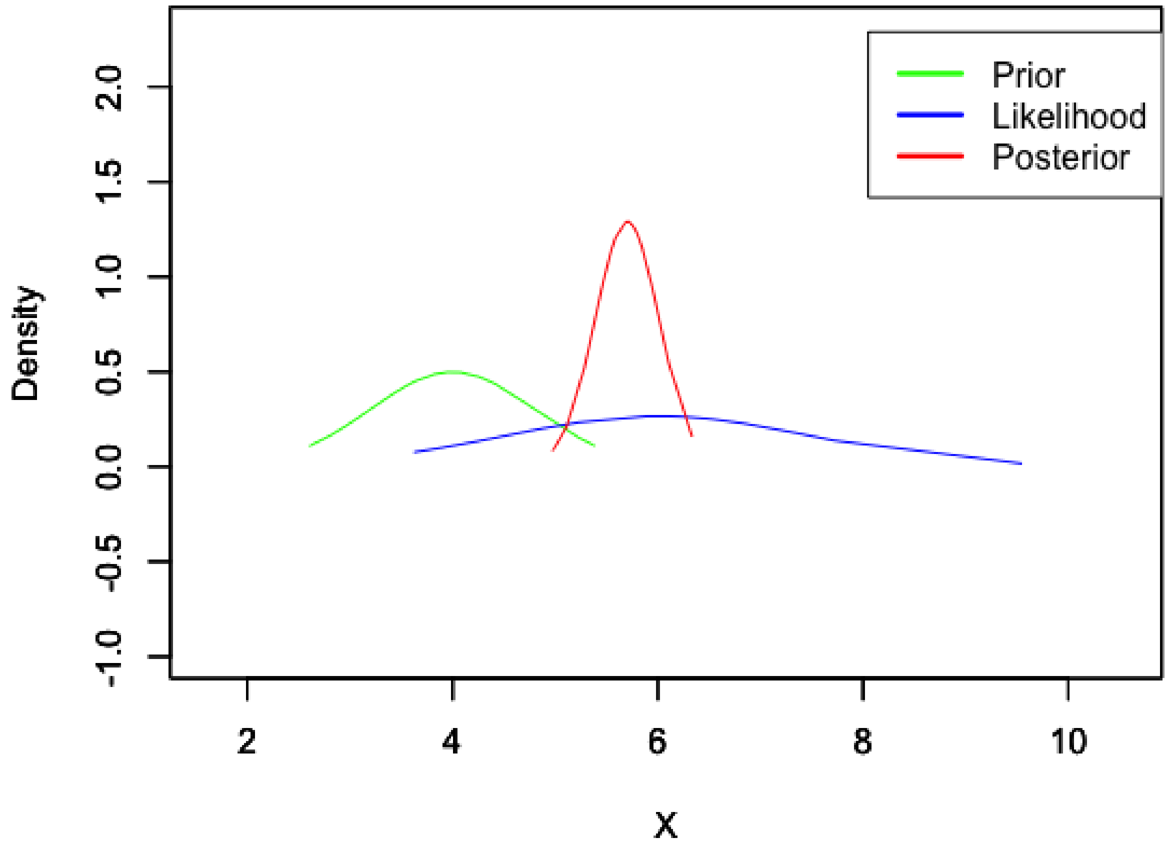
$$\sigma_n^2 = \left(\frac{\sigma^2\sigma_0^2}{\sigma^2 + n\sigma_0^2}\right) = \frac{1.5^2 0.8^2}{(1.5^2 + 20(0.8^2))} = \frac{1.44}{15.05} = 0.09568106$$

$$\mu_n = \left(\frac{n\bar{X}\sigma_0^2 + \mu_0\sigma^2}{\sigma^2 + n\sigma_0^2}\right) = \frac{20(6)(0.8^2) + 4(1.5^2)}{(1.5^2 + 20(0.8^2))} = \frac{85.8}{15.05} = 5.70099668$$

These are the values for the parameters of posterior distribution.

Using these the plots for all three distributions were plotted.
It is included below:

Also, the R code for the same is also included with the submission.



Prior
Likelihood
Posterior
Density
X
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
2
4
6
8
10