Name: Rishal Shah UnityID: rshah27

Bayesian Parameter Estimation

Q1:

$$Since,$$

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

$$p(\mu|X) \propto p(\mu) * \prod_{i=1}^{n} p(x_i|\mu)$$

$$= \frac{1}{\sqrt{2\pi\sigma_0^2}} * e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}} * \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} * e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$= \frac{1}{\sqrt{2\pi}^{n+1} * \sqrt{\sigma_0^2 * \sigma^{2n}}} * e^{\frac{-\mu^2+2\mu\mu_0-\mu_0^2}{2\sigma_0^2} - \sum_{i=1}^{n}\frac{x_i^2-2\mu x_i+\mu^2}{2\sigma^2}}$$

$$\propto e^{\left[\frac{-\mu^2\sigma^2+2\mu\mu_0\sigma^2-\mu_0^2\sigma^2-\sigma_0^2\sum x_{\mathrm{i}}^2+2\mu\sigma_0^2\sum x_{\mathrm{i}}-n\mu^2\sigma_0^2}{2\sigma_0^2\sigma^2}\right]}$$

$$= e^{\left[\frac{-\mu^2(\sigma^2+n\sigma_0^2)+2\mu(\mu_0\sigma^2+\sigma_0^2\sum x_{\mathrm{i}})-(\mu_0^2\sigma^2+\sigma_0^2\sum x_{\mathrm{i}}^2)}{2\sigma_0^2\sigma^2}\right]}$$

$$p(\mu|X)$$

$$= e^{\left[\frac{-\mu^2+\frac{2\mu(\mu_0\sigma^2+\sigma_0^2\sum x_{\mathrm{i}})}{\sigma^2+n\sigma_0^2}-\left(\frac{\mu_0\sigma^2+\sigma_0^2\sum x_{\mathrm{i}}}{\sigma^2+n\sigma_0^2}\right)^2+\left(\frac{\mu_0\sigma^2+\sigma_0^2\sum x_{\mathrm{i}}}{\sigma^2+n\sigma_0^2}\right)^2}{\frac{2\sigma_0^2\sigma^2}{\sigma^2+n\sigma_0^2}}\right]} * e^{-\frac{(\mu_0^2\sigma^2+\sigma_0^2\sum x_{\mathrm{i}}^2)}{2\sigma_0^2\sigma^2}}$$

$$\propto e^{\left[\frac{-\mu^2+\frac{2\mu(\mu_0\sigma^2+\sigma_0^2\sum x_{\mathrm{i}})}{\sigma^2+n\sigma_0^2}-\left(\frac{\mu_0\sigma^2+\sigma_0^2\sum x_{\mathrm{i}}}{\sigma^2+n\sigma_0^2}\right)^2}{\frac{2\sigma_0^2\sigma^2}{\sigma^2+n\sigma_0^2}}\right]}$$

$$\therefore \boxed{p(\mu|X) \propto e^{\left[\frac{-\left(\mu-\left(\frac{\mu_0\sigma^2+\sigma_0^2\sum x_{\mathrm{i}}}{\sigma^2+n\sigma_0^2}\right)\right)^2}{\frac{2\sigma_0^2\sigma^2}{\sigma^2+n\sigma_0^2}}\right]}} \quad ----(\mathrm{A})$$

Q2:

From equation (A), we can see that the posterior distribution is quadratic and a Gaussian distribution as it is proportional to a Normal Distribution of the form:

$$e^{-\frac{(\mu-\mu_n)^2}{2\sigma_n^2}} \sim N(\mu_n, \sigma_n^2)$$

$$\text{where, } \sigma_{\text{n}}^2 = \frac{\sigma_0^2\sigma^2}{\sigma^2+n\sigma_0^2} \ and \ \mu_n = \frac{\mu_0\sigma^2+\sigma_0^2\sum x_{\text{i}}}{\sigma^2+n\sigma_0^2}$$

Q3:

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

$$\therefore \boxed{\frac{1}{\sigma_n^2} = \frac{1}{\sigma_0^2}+\frac{n}{\sigma^2}} ----\text{(B)}$$

$$\mu_n = \frac{\mu_0\sigma^2+\sigma_0^2\sum x_{\text{i}}}{\sigma^2+n\sigma_0^2}$$

$$= \frac{\mu_0\sigma^2}{\sigma^2+n\sigma_0^2} + \frac{\sigma_0^2\sum x_{\text{i}}}{\sigma^2+n\sigma_0^2}$$

$$= \frac{\mu_0\sigma^2}{\sigma^2+n\sigma_0^2} + \frac{\sigma_0^2 * n * \frac{\sum x_{\text{i}}}{n}}{\sigma^2+n\sigma_0^2}$$

$$\boxed{\mu_n = \frac{\mu_0\sigma^2}{\sigma^2+n\sigma_0^2} + \frac{n\bar{\text{x}}\sigma_0^2}{\sigma^2+n\sigma_0^2}} ----\text{(C)}$$

Q4:

From eq. (C) we can identify the weights:

$$\boxed{w_0 = \frac{\sigma^2}{\sigma^2+n\sigma_0^2} \ and \ w_1 = \frac{n\sigma_0^2}{\sigma^2+n\sigma_0^2}} ----\text{(D)}$$

Q5:

From equation 4, we can see that weight w_0 and w_1, are **inversely proportional** to their variances.

Q6:

Sum of weights =

$$w_0 + w_1 = \frac{\sigma^2}{\sigma^2 + n\sigma_0^2} + \frac{n\sigma_0^2}{\sigma^2 + n\sigma_0^2}$$

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

$$\boxed{\therefore w_0 + w_1 = 1}$$

Q7:

Since the weights $w_0 = \frac{\sigma^2}{\sigma^2+n\sigma_0^2} = \frac{1}{1+\frac{n\sigma_0^2}{\sigma^2}}$, here numerator = 1 and denominator is greater than 1. So w_0 is between 0 and 1.

$and\ w_1 = \frac{n\sigma_0^2}{\sigma^2+n\sigma_0^2} = \frac{1}{1+\frac{\sigma^2}{n\sigma_0^2}}$ also has numerator = 1 and denominator greater than 1. So w_1 is also between 0 and 1.

Therefore, both the weights are between 0 and 1.

Q8:

From the previous answers 4-7, we can infer that, since both weights have their value between 0 and 1, the maximum value of μ_n is $\mu_0 + \bar{x}$ (when both weights are 1) and 0 (when both weights are 0). But, since both weights are inversely dependent on n, the edge cases of 0 and 1 are not possible. Hence, μ_n will always lie between μ_0 and $\bar{x}$.

Q9:

$$p(x_{new}|X) = \int p(x_{new}|\mu)\, p(\mu|X)\, d\mu$$

Since,

$$x_{new} = (x_{new} - \mu) + \mu,$$

where, $x_{new} - \mu \sim N(0, \sigma^2)$ (Normal distribution)

and $\mu \sim N(\mu_n, \sigma_n^2)$ (Normal distribution)

$x_{new} = (x_{new} - \mu) + \mu$

$\mathrm{p}(x_{new}|\mathrm{X}) \sim N(0, \sigma^2) + N(\mu_n, \sigma_n^2)$

This is because if there are two mutually independent normal random variables with means μ_1, μ_2 and variances σ_1^2, σ_2^2, then the linear combination: $Y = \sum_{i=1}^{n} c_i X_i$ follows the normal distribution:

$$N(\sum_{i=1}^{n} c_i \mu_i\, , \sum_{i=1}^{n} c_i^2 \sigma_i^2)$$

Hence,

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

Q10: Code (Python3):

```
import numpy as np
import scipy.stats as st
from matplotlib import pyplot as plt

sample = 20
line = np.linspace(0, 10, sample)

m0 = 4
sd0 = 0.8
d_prior = st.norm(m0, sd0).pdf(line)

mx = 6
sdx = 1.5
```

```
d_sample = st.norm(mx, sdx).pdf(line)

x_t = st.norm(mx,sdx).rvs(sample)
variance = 1/((1/sd0**2)+(sample/sdx**2))
mean = variance * ((m0/sd0**2)+(np.mean(x_t)*sample/sdx**2)
)
print("Posterior distibution:\nMean =",round(mean,3))
print("Variance =",round(variance,3))
d_posterior = st.norm(mean, np.sqrt(variance)).pdf(line)

plt.figure(figsize=(20,20))
plt.plot(line,d_prior,"r-",label='Prior Distribution')
plt.plot(line,d_sample,"g-",label='Sample Distribution')
plt.plot(line,d_posterior,"b-
",label='Posterior Distribution')
plt.legend(loc='upper right')
plt.title('Probability Density Plot')
plt.ylabel('Probability Density')
plt.xlabel('X')
plt.show()
```

Posterior distibution:
Mean = 5.727
Variance = 0.096