Bayesian Parameter Estimation: CSC 591 ADBI

Work Done by: Ashwin Balasubramaniam (abalasu4@ncsu.edu)

Given:

- **Univariate Case: The data** 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,** p(x)-N(μ,σ²)**, with parameters** μ² **and** σ²**.**

$$p(\mathrm{x}|\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(\mu-\mu\ \ )^2}{2\sigma^2}}$$

- **Parameters: Unknown mean, known variance**
- **Priors: The conjugate prior for** μ **is Gaussian,** p(μ)~N(μ_0,σ_0^2)

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

1. **Derive the formula for the posterior distribution of μ**

$$p(\mu|X_t) = p(\mu)p(X_t|\mu)$$

Where $p(\mu|X_t) - posterior\ distribution$

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

$$= e^{\frac{-\mu^2(\sigma^2+n\sigma_0^2)+2\mu(\mu_0\sigma^2+\sigma_0^2\sum x_t)-(\mu_0^2\sigma^2+\sigma_0^2\sum x_t^2)}{2\sigma_0^2\sigma^2}}$$

$$p(\mu|X) = e^{\frac{-\mu^2+\frac{2\mu(\mu_0\sigma^2+\sigma_0^2\sum x_t)}{\sigma^2+n\sigma_0^2}-\left(\frac{\mu_0\sigma^2+\sigma_0^2\sum x_t}{\sigma^2+n\sigma_0^2}\right)^2+\left(\frac{\mu_0\sigma^2+\sigma_0^2\sum x_t}{\sigma^2+n\sigma_0^2}\right)^2}{\frac{2\sigma_0^2\sigma^2}{\sigma^2+n\sigma_0^2}}} * e^{-\frac{(\mu_0^2\sigma^2+\sigma_0^2\sum x_t^2)}{2\sigma_0^2\sigma^2}}$$

$$p(\mu|X) \propto e^{\frac{-\mu^2+\frac{2\mu(\mu_0\sigma^2+\sigma_0^2\sum x_t)}{\sigma^2+n\sigma_0^2}-\left(\frac{\mu_0\sigma^2+\sigma_0^2\sum x_t}{\sigma^2+n\sigma_0^2}\right)^2}{\frac{2\sigma_0^2\sigma^2}{\sigma^2+n\sigma_0^2}}}$$

$$p(\mu|X) \propto e^{\frac{-\left(\mu-\left(\frac{\mu_0\sigma^2+\sigma_0^2\sum x_t}{\sigma^2+n\sigma_0^2}\right)\right)^2}{\frac{2\sigma_0^2\sigma^2}{\sigma^2+n\sigma_0^2}}}$$

2. **Show that the posterior distribution is Gaussian , p(μ|X)~N(μn, σn)**

   From the previous questions it can be seen that the posterior distribution is a Gaussian distribution since its proportional to Normal Distribution

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

$$Here\ , \sigma_n = \frac{\sigma_0^2\sigma^2}{\sigma^2 + n\sigma_0^2}$$

$$\mu_n = \frac{\mu_0\sigma^2 + \sigma_0^2\sum x_t}{\sigma^2 + n\sigma_0^2}$$

3. **Show the derivation and final estimate for μn, 1/σ²n**

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

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

$$\mu_n = \frac{\mu_0\sigma^2 + \sigma_0^2\sum x_t}{\sigma^2 + n\sigma_0^2} = \frac{\frac{\mu_0\sigma^2}{\sigma_0^2\sigma^2} + \frac{\sigma_0^2\sum x_t}{\sigma_0^2\sigma^2}}{\frac{\sigma^2 + n\sigma_0^2}{\sigma_0^2\sigma^2}}$$

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

4. **If the mean of the posterior density (which is the MAP estimate), μn is written as the weighted average of the prior mean, μ0, and the sample (likelihood) mean, $\bar{X}$, then what are the formulas for the weights?**

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

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

$$= w_0 * \mu_0 + w_1 * \bar{x}$$

$$w_0 = \frac{\sigma_n^2}{\sigma_0^2} \; and \; w_1 = \frac{n\sigma_n^2}{\sigma^2}$$

$$where \; w_0 \; and \; w_1 are \; the \; weights$$

5. **Are the weights in Question #4 directly or inversely proportional to their variances (justify)?**

From the previous question it can be seen that the weights are inversely proportional to their variances. W_0 corresponds to prior mean and it is inversely proportional to its variance. W_1 is the weighted average of likelihood and its inversely proportional to its variance. Hence, both the weights are inversely proportional to their variances

6. **Do the weights in Questions #4 sum up to 1 (justify)?**

Yes, the weights sum up to 1.

$$w_0 + w_1 = \frac{\sigma_n^2}{\sigma_0^2} + \frac{n\sigma_n^2}{\sigma^2}$$
$$= \frac{1}{\sigma_0^2 * \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}\right)} + \frac{n}{\sigma^2 * \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}\right)}$$
$$\frac{\sigma^2}{\sigma^2 + n\sigma_0^2} + \frac{n\sigma_0^2}{\sigma^2 + n\sigma_0^2} = 1$$

7. **Is each weight between zero and one (justify)?**

From the previous question it can be seen that the sum of weights is 1 i.e

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

Hence, it can be concluded that each weight will be between 0 and 1

In w_0 and w_1 it can be seen that the denominator has two terms and will always be greater than the numerator. This is because, variance σ^2 cannot be negative and the n cannot be negative.
Hence, it can be concluded that each weight will be between 0 and 1

8. **Given your answers for Questions #4-7, what can you say about the value of μ_n w.r.t. the values of μ_0 and $\overline{X}$.**

We know that $\mu_n = w_0 * \mu_0 + w_1 * \bar{x}$

The sum of weights $w_0 + w_1 = 1$.

Both the weights depend on n. When n increases , the weights vary inversely to each other. It can be concluded that the values of μ_n lie between μ_0 and $\bar{x}$.

9. **If** σ^2 **is known, then for the new instance** x^{new}**, show that** $p(x^{new}|X) \sim N(\mu_n, \sigma_n^2+\sigma^2)$

$p(x_i|\mu)$ is given to be a normal distribution $N(\mu, \sigma^2)$ and we found the posterior distribution $p(\mu|X)$ to also be a normal distribution $N(\mu_n, \sigma_n^2)$.

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

Here, μ is the normal distribution N(0, σ^2) and μ is the posterior normal distribution N($\mu_{n\,,}$ σ_2)

WKT , sum of two normal distributions is a normal distribution,
Therefore,

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

$$\sim N(0, \sigma^2) + N(\mu_n, \sigma_n^2)$$

$$\sim N(\mu_n, \sigma^2 + \sigma_n^2)$$

10. **Generate a plot that displays** $p(x) \sim N(6,1.5^2)$, **prior** $p(\mu) \sim N(4,0.8^2)$ **and posterior** $p(\mu|X) \sim N(\mu_n, \sigma_n^2)$ **for** n=20 **sample points. What are the values for** μ_n **and** σ_n^2**?**

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

s = 20
d = np.linspace(0,10, s)
meanp = 4
stdp = 0.82
normprior = st.norm(meanp,stdp).pdf(d) #prior distribution


meanx = 6
stdx = 1.5
normx = st.norm(meanx,stdx).pdf(d) #sample distribution


t = st.norm(meanx,stdx).rvs(s)
var = 1/((1/stdp**2) + (s/stdx ** 2))
mean = var * ((meanp/stdp**2) + (np.mean(t)*s/stdx**2))
normpost = st.norm(mean,np.sqrt(var)).pdf(d) #posterior distribution

print("Mean : ",mean ,"Variance : ",var)
```

Mean : 5.813920458805838 Variance : 0.096375334437507 97

```
plt.plot(d,normprior,color = 'red' , label = 'Prior Distribution')
plt.plot(d,normx,color = 'blue', label = 'Sample Distribution')
plt.plot(d,normpost , color = 'green',label = 'Posterior Distribution')
plt.legend(loc = 'upper left')
plt.xlabel ('X')
plt.title('Probability Density Plot')
plt.show()
plt.savefig('pdf.png')
```