Algorithms for Data Guided Business Intelligence

Home Work

Topic 2, Part 2

By Rohit Mandge

Unity ID: ***RNMANDGE***

Q.1] $P(\mu \mid data) = P(data \mid \mu) * P(\mu)$

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

$$\Rightarrow P(data \mid \mu) = P(x^1 \mid \mu) \cdot P(x^2 \mid \mu) \cdots P(x^n \mid \mu)$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} * \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x_2-\mu)^2}{2\sigma^2}}$$

$$* \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x_3-\mu)^2}{2\sigma^2}} * \cdots \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x_n-\mu)^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot e^{-\frac{\sum_{t=1}^{n}(x_t-\mu)^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot e^{-\frac{\sum_{t=1}^{n}(x_t^2+\mu^2-2x_t\mu)}{2\sigma^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot e^{\frac{-n\mu^2 + \sum_{t=1}^{n}(x_t^2-2x_t\mu)}{2\sigma^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot e^{\frac{-n\mu^2+2\mu\cdot\sum_{t=1}^{n}x_t - \sum_{t=1}^{n}x_t^2}{2\sigma^2}}$$

$$P(\mu \mid data) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot \left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right) \cdot e^{\left(\frac{-n\mu^2+2\mu\cdot\sum_{t=1}^{n}x_t - \sum_{t=1}^{n}x_t^2}{2\sigma^2} - \frac{(\mu-\mu_0)^2}{2\sigma_0^2}\right)}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right) \cdot e^{-\left[\frac{\sigma_0^2 (n\mu^2 - 2\mu\sum_{t=1}^{n} x_t + \sum_{t=1}^{n} x_t^2) + \sigma^2(\mu-\mu_0)^2}{2\sigma^2 \cdot \sigma_0^2}\right]}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right) \cdot e^{-\left[\frac{\sigma_0^2 n\mu^2 - \sigma_0^2(2\mu\sum_{t=1}^{n} x_t + \sum_{t=1}^{n} x_t^2) + \sigma^2\mu^2 - 2\sigma^2\mu\mu_0 + \sigma^2\mu_0^2}{2\sigma^2\sigma_0^2}\right]}$$

$$\boxed{\left\{= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot \left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right) \cdot e^{-\left[\frac{\sigma_0^2 n\mu^2 - \sigma_0^2(2\mu\sum_{t=1}^{n} x_t + \sum_{t=1}^{n} x_t^2) + \sigma^2\mu^2 - 2\sigma^2\mu\mu_0 + \sigma^2\mu_0^2}{2\sigma^2\sigma_0^2}\right]}\right\}}$$

Q. 2]

→ proving the posterior distribution is Gaussian

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot \left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right) \cdot e^{-\left[\frac{\sigma_0^2 n \cdot \mu^2 - \sigma_0^2(2\mu\sum_{t=1}^{n} x_t + \sum_{t=1}^{n} x_t^2) + \sigma^2\mu^2 - 2\sigma^2\mu\mu_0 + \sigma^2\mu_0^2}{2\sigma^2\sigma_0^2}\right]} \quad \text{— (1)}$$

comparing with,

$$p(\mu|X) \sim N(\mu_n, \sigma_n^2)$$

$$= \frac{1}{\sqrt{2\pi\sigma_n^2}} \cdot e^{\frac{-(\mu_n - \mu)^2}{2\sigma_n^2}}$$

$$= \frac{1}{\sqrt{2\pi\sigma_n^2}} \cdot e^{\frac{-(\mu_n^2 + \mu^2 - 2\mu_n\mu)}{2\sigma_n^2}} \quad \text{— (2)}$$

→ taking co-efficients of eq^n. (1)

$$\left[= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot \left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right) \cdot e^{-\left[\frac{(\sigma_0^2 n + \sigma^2)\mu^2 - (2\sigma_0^2\sum_{t=1}^{n} x_t + 2\sigma^2\mu_0)\mu + (\sigma_0^2\sum_{t=1}^{n} x_t^2 + \sigma^2\mu_0^2)}{2\sigma^2\sigma_0^2}\right]}\right.$$

Hence, the equation is in Gaussian form.

Q. 3] Derivation & the final estimate μ_n & $1/\sigma_n^2$.

→ comparing equations from Q. 2],

we get,

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

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

Now, we find, $\underline{\mu_n}$, -

$$-\frac{\cancel{2}\cancel{\mu}\mu_n}{2\sigma_n^2} = -\cancel{2}\frac{\left[\sigma_0^2 \sum_{t=1}^{n} x_t + \sigma^2\mu_0\right]\cancel{\mu}}{2\sigma^2\sigma_0^2}.$$

$$\frac{\mu_n}{\cancel{2}\sigma_n^2} = \frac{\sigma_0^2 \sum_{t=1}^{n} x_t + \sigma^2\mu_0}{\cancel{2}\sigma^2\sigma_0^2}$$

$$\mu_n = \sigma_n^2 \left[\frac{\sigma_0^2 \sum_{t=1}^{n} x_t + \sigma^2\mu_0}{\sigma^2\sigma_0^2}\right]$$

$$= \frac{\cancel{\sigma^2\sigma_0^2}}{n\sigma_0^2+\sigma^2}\left[\frac{\sigma_0^2 \sum_{t=1}^{n} x_t + \sigma^2\mu_0}{\cancel{\sigma^2\sigma_0^2}}\right]$$

$$= \boxed{\frac{\sigma_0^2 \sum_{t=1}^{n} x_t + \sigma^2\mu_0}{n\sigma_0^2+\sigma^2}}$$

Q.4] To get the weigted Avg,

$$\mu_n = \frac{\sigma_0^2 \sum_{t=1}^{n} x_t + \sigma^2 \mu_0}{n\sigma_0^2 + \sigma^2}$$

$$= \frac{\sigma_0^2 n . \bar{x} + \sigma^2 \mu_0}{n\sigma_0^2 + \sigma^2}$$

$$= \underbrace{\left(\frac{\sigma_0^2 . n}{n\sigma_0^2 + \sigma^2}\right)}_{w_1} \bar{x} + \underbrace{\left(\frac{\sigma^2}{n\sigma_0^2 + \sigma^2}\right)}_{w_2} \mu_0 .$$

It can be expressed in weighted avg. form.

Q.5]

the weights found in Q.4 (above) are inversely proportional to their respective variances.

As, 'σ^2' is in the denominator, the more the value of 'σ', less would be the value of ~~variance~~ weights.
hence, they are inversely proportional.

{∵ variance of μ_0 is σ_0 & variance of $\bar{x}$ is 'σ'.}

Q.6] from Q.4, we get;

$$w_1 = \frac{\sigma_0^2 n}{\sigma_0^2 n + \sigma^2} \quad — ①$$

$$w_2 = \frac{\sigma^2}{\sigma_0^2 n + \sigma^2} \quad — ②$$

Adding eq^n ① & ②, we get,

$$\left\{ w_1 + w_2 = \frac{\sigma_0^2 n + \sigma^2}{\sigma_0^2 n + \sigma_0^2} = \boxed{1} \right\}$$

Q.7] As, σ, σ_0 & n are positive,

$$w_1 = \frac{\sigma_0^2 n}{\sigma_0^2 n + \sigma^2}$$ are between 0 & 1

$$\because \sigma_0^2 n + \sigma^2 \geq \sigma_0^2 n$$

$$w_2 = \frac{\sigma^2}{\sigma_0^2 n + \sigma^2}, \quad \because \sigma_0^2 n + \sigma^2 \geq \sigma^2$$

hence,

$$\therefore \boxed{0 < w_1, w_2 < 1}$$

Also, as mentioned in (Q.6) $w_1 + w_2 = 1$,

hence, $\left[0 < w_1, w_2 \leq 1\right]$

Q.8], from Q.3 & Q.4

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

~~Q.10~~]

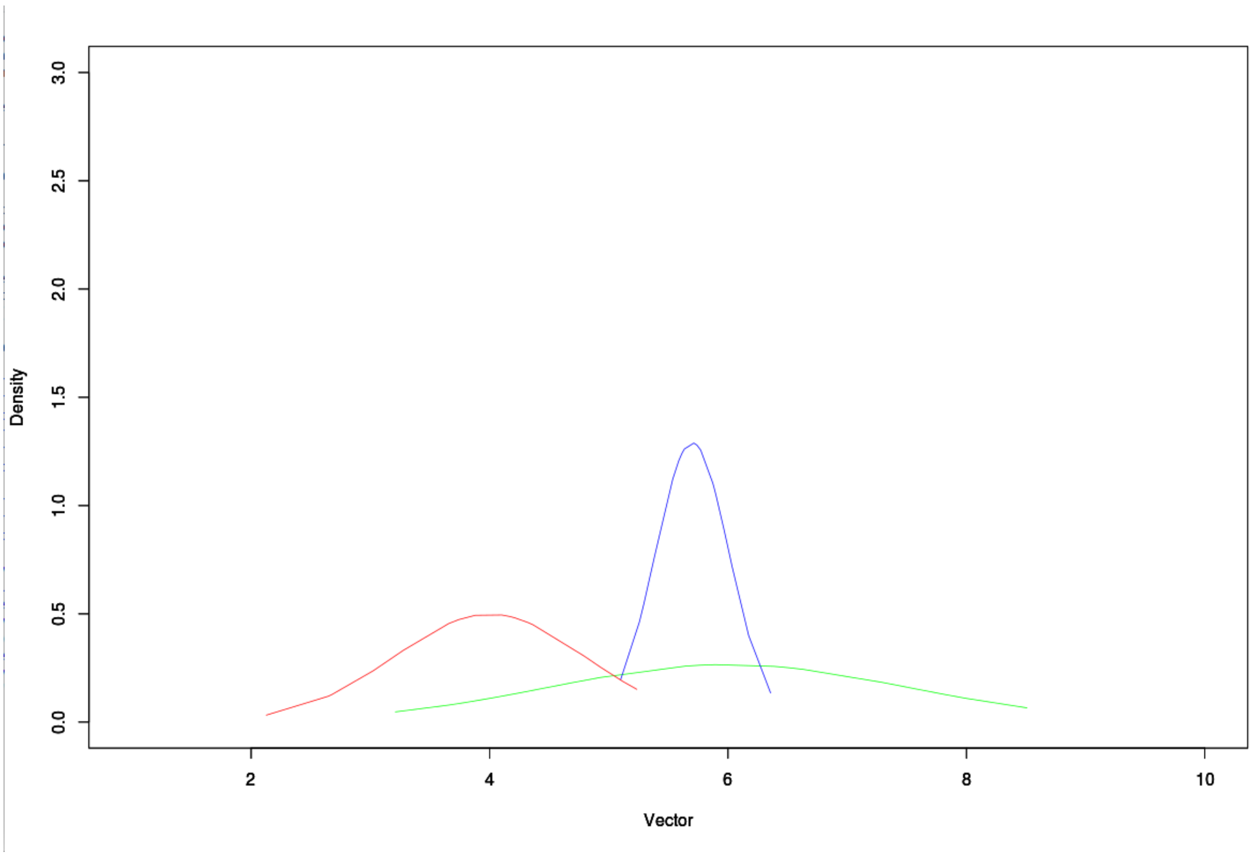
Q.10)

```
vectorPrior <- rnorm(20, mean = 4, sd = 0.8)
vectorPrior <- sort(vectorPrior)
densityPrior <- dnorm(vectorPrior, mean = 4, sd = 0.8)
vectorLikelihood <- rnorm(20, mean = 6, sd = 1.5)
vectorLikelihood <- sort(vectorLikelihood)
densityLikelihood <- dnorm(vectorLikelihood, mean = 6, sd = 1.5)

vectorPost <- rnorm(20, mean = 5.7, sd = 0.3093236)
vectorPost <- sort(vectorPost)
densityPost <- dnorm(vectorPost, mean = 5.7, sd = 0.3093236)

plot(vectorLikelihood, densityLikelihood, type = "l", col="green",
xlim=c(1,10), ylim = c(0,3), xlab="Vector", ylab="Density")
par(new=TRUE)
plot(vectorPost, densityPost, type = "l", col="blue", xlim=c(1,10), ylim =
c(0,3), xlab="Vector", ylab="Density")
par(new=TRUE)
plot(vectorPrior, densityPrior, type = "l", col="red", xlim=c(1,10), ylim =
c(0,3), xlab="Vector", ylab="Density")
```

The values we get are mean = 5.7 and variance = (0.3)^2


Density
0.0
0.5
1.0
1.5
2.0
2.5
3.0
2
4
6
8
10
Vector