MAP Estimate:

$$D = \{x_1, x_2, \ldots, x_n\}$$

$$p_{\theta|D} = \frac{p_{D|\theta} \; p_\theta}{Z}$$

Consider estimating the mean of Gaussian density with variance assumed known.

$$p_{x|\mu} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ \frac{-1}{2\sigma^2} (x-\mu)^2 \right\}$$

μ is the only parameter.

$$\mu_{MLE} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$p_{\mu|D} = \frac{p_{D|\mu} \; p_\mu}{Z}$$

- For conjugate prior, we want p_μ & $p_{\mu|D}$ to have the same fun^c form

$$p_{D|\mu} = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left\{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$$

As a fun^c of μ, this has an exponential of a quadratic in μ

* we need p_μ s.t. when multiplied by $p_{D|\mu}$ we need to get the same form as of p_μ

* $\therefore$ the conjugate prior here is a normal density

* $p_\mu = \mathcal{N}(\mu_0, \sigma_0^2)$

where μ_0, σ_0^2 are hyper parameters.

$$p_{\mu/D} = \frac{p_{D|\mu}\; p_\mu}{Z}$$

$$p_{\mu/D} \propto \exp\left\{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 - \frac{1}{2\sigma_0^2}(\mu-\mu_0)^2\right\}$$

$$= \exp\left\{-\frac{1}{2}A\right\}$$

$$A = \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i^2 + \mu^2\left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}\right)$$

$$- 2\mu\left(\sum_{i=1}^{n}\frac{x_i}{\sigma^2} + \frac{\mu_0}{\sigma_0^2}\right) + \frac{\mu_0^2}{\sigma_0^2} \rightarrow (1)$$

Suppose we have $p_{\mu|D}$ is $N(\mu_n, \sigma_n^2)$

$$p_{\mu|D} \propto \exp\left\{-\frac{1}{2}\frac{(\mu-\mu_n)^2}{\sigma_n^2}\right\}$$

$$= \exp\left\{-\frac{1}{2}\left[\frac{\mu^2}{\sigma_n^2} + \frac{\mu_n^2}{\sigma_n^2} - 2\mu\frac{\mu_n}{\sigma_n^2}\right]\right\}$$

$\rightarrow (2)$

now Comparing (1) & (2)

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

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

$$\frac{\mu_n}{\sigma_n^2} = \frac{1}{\sigma^2} \sum_{i=1}^{n} x_i + \frac{\mu_0}{\sigma_0^2}$$

$$\mu_n = \frac{\sigma_n^2}{\sigma^2} \sum_{i=1}^{n} x_i + \frac{\sigma_n^2}{\sigma_0^2} \mu_0$$

μ_n is a convex combⁿ of MLE & μ_0

* for large n $\mu_n \approx \mu_{MLE}$

As n becomes very large Bayesian Estimate is same as MLE

* Since the distribution is Gaussian mode as well as mean is μ_n