4/12/21 Lec16

$$\hat{\theta}_{MMSE} = \frac{\frac{n\bar{x}}{\sigma^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}} + \frac{\frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}$$

$$= \frac{\frac{n}{\sigma^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}\hat{\theta}_{MLE}\cdot\frac{\frac{\sigma^2}{n}}{\frac{\sigma^2}{n}} + \frac{\tau^2}{\tau^2}\underbrace{\frac{\frac{1}{\tau^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}}_{\rho}E[\theta|\sigma^2]$$

$$= \frac{1}{1+\frac{\sigma^2}{n\tau^2}}\hat{\theta}_{MLE}\frac{n\tau^2}{n\tau^2} + \frac{\sigma^2}{\sigma^2}E[\theta|\sigma^2]$$

$$= \underbrace{\frac{n\tau^2}{n\tau^2+\sigma^2}}_{1-\rho}\hat{\theta}_{MLE} + \underbrace{\frac{\sigma^2}{n\tau^2+\sigma^2}}_{\rho}E[\theta|\sigma^2]$$

What is Laplace's Prior? $p(\theta|\sigma^2) \propto 1$ Lec 15

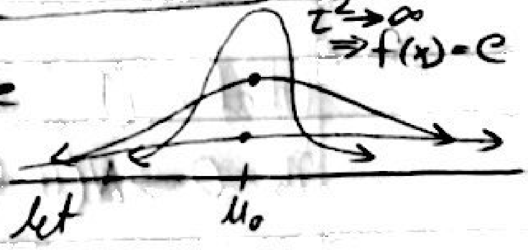
$$p(\theta|X,\sigma^2) \propto p(X|\theta,\sigma^2)p(\theta|\sigma^2) \propto p(X|\theta,\sigma^2) \propto N\left(\bar{x},\frac{\sigma^2}{n}\right)$$

$$= N\left(\frac{\frac{n\bar{x}}{\sigma^2}+\frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}, \frac{1}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}\right) \Rightarrow \frac{\sigma^2}{n} = \frac{1}{\frac{n}{\sigma^2}+\cancel{\frac{1}{\tau^2}}} \Rightarrow \tau^2=\infty$$

$$\frac{\frac{n\bar{x}}{\sigma^2}+\frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}} = \bar{X} \Rightarrow \mu_0 \text{ could be any value...}$$

by convention we will let it equal to zero.

$\tau^2 \to \infty \Rightarrow f(x) = c$



$\Rightarrow p(\theta|\sigma^2) = N(0,\infty)$ is Laplace's prior (improper)

But is the posterior proper? Yes, Always!

same as Laplace.

Jeffrey's Prior: $p_J(\theta|\sigma^2) \propto \sqrt{I(\theta|\sigma^2)} = \sqrt{\frac{n}{\sigma^2}} \propto 1 \propto N(0,\infty)$

$$l'(\theta;x,\sigma^2) = \frac{n\bar{x}}{\sigma^2} - \frac{n\theta}{\sigma^2} \Rightarrow l''(\theta;x,\sigma^2) = \frac{-n}{\sigma^2}$$

$$\Rightarrow I(\theta|\sigma^2) = E_x[-l''(\theta;x,\sigma^2)] = E_x\left[\frac{n}{\sigma^2}\right] = \frac{n}{\sigma^2}$$

Scanned with CamScanner

We want a pseudocount interpretation of the hyperparameters μ_0 and τ^2. The best way to do this, is to do a small reparameterization of the prior's τ^2. Recall that we know σ^2.

$$\tau^2 := \frac{\sigma^2}{n_0} \Rightarrow P(\theta|X,\sigma^2) = N\left(\frac{\frac{n\bar{X}}{\sigma^2}+\frac{n_0\mu_0}{\sigma^2}}{\frac{n}{\sigma^2}+\frac{n_0}{\sigma^2}}, \frac{1}{\frac{n}{\sigma^2}+\frac{n_0}{\sigma^2}}\right) \text{ multiply } \sigma^2$$

$$= N\left(\frac{\underbrace{n\bar{X}}_{\Sigma X_i}+\underbrace{n_0\mu_0}_{\Sigma Y_i}}{n+n_0}, \frac{\sigma^2}{n+n_0}\right)$$

So n_0 represents number of pseudo-observations. What does μ_0 represents?

Let $Y_1, Y_2, \dots Y_{n_0}$ be the "pseudodata". Let $\mu_0 = \bar{Y}$, be the sample average of the pseudodata and $n_0\mu_0$ is the sum of the pseudodata.

$$\Rightarrow \theta_{MMSE} = \frac{n}{n+n_0}\bar{X} + \overbrace{\frac{n_0}{n+n_0}}^{\rho}\mu_0$$

What's the Haldane prior of total ignorance?

$$n_0 = 0 \Rightarrow N\left(\mu_0, \frac{\sigma^2}{n_0}\right) = N(0, \infty)$$

This means all three objective priors we studied are the same.

What is the posterior predictive distribution for $n_* = 1$ observ?

$$P(X_*|X,\sigma^2) = \int_{\Theta} \underbrace{P(X_*|\theta,\sigma^2)}_{N(\theta,\sigma^2)} \underbrace{P(\theta|X,\sigma^2)}_{N(\theta_p,\sigma_p^2)} d\theta$$

$$= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(X_*-\theta)^2} \cdot \frac{1}{\sqrt{2\pi\sigma_p^2}} e^{-\frac{1}{2\sigma_p^2}(\theta-\theta_p)^2} d\theta$$

Scanned with CamScanner

$$\propto \int_{\mathbb{R}} e^{\frac{-X_r^2}{2\sigma^2}} e^{\frac{X_r\theta}{\sigma^2}} e^{-\frac{\theta^2}{2\sigma^2}} \cdot e^{\frac{-\theta^2}{2\sigma_p^2}} e^{\frac{\theta\theta_p}{\sigma_p^2}} e^{\frac{-\theta_p^2}{2\sigma_p^2}} \, d\theta$$

$$\propto e^{\frac{-X_r^2}{2\sigma^2}} \int_{\mathbb{R}} \overbrace{e^{a\theta - b\theta^2}}^{\propto \text{Normal}} \, d\theta \; ; \quad a = \frac{X_r}{\sigma^2} + \frac{\theta_p}{\sigma_p^2}$$

$$b = +\frac{1}{2}\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_p^2}\right)$$

$$P(\theta) = N\left(\frac{a}{2b}, \frac{1}{2b}\right) = \frac{1}{\sqrt{2\pi\left(\frac{1}{2b}\right)}} e^{-\frac{1}{2(1/2b)}\left(\theta - \frac{a}{2b}\right)^2}$$

$$= \sqrt{\frac{b}{\pi}} \, e^{-b\left(\theta^2 - \frac{\theta a}{b} + \frac{a^2}{4b^2}\right)}$$

$$= \sqrt{\frac{b}{\pi}} \, e^{-b\theta^2 + \theta a - \frac{a^2}{4b}}$$

$$= \underbrace{\sqrt{\frac{b}{\pi}} \, e^{\frac{-a^2}{4b}}}_{c} e^{a\theta - b\theta^2}$$

$$= e^{\frac{-X_r^2}{2\sigma^2}} \frac{1}{c} \overbrace{\int_{\mathbb{R}} c \, e^{-a\theta - b\theta^2} \, d\theta}^{1}$$

$$= e^{\frac{-X_r^2}{2\sigma^2}} \sqrt{\frac{\pi}{b}} \, e^{\frac{a^2}{4b}}$$

$$\propto e^{\frac{-X_r^2}{2\sigma^2}} \left(\frac{1}{2}\left(\frac{1}{\sigma} + \frac{1}{\sigma_p^2}\right)\right)^{-\frac{1}{2}} e^{\frac{\left(\frac{X_r}{\sigma^2} + \frac{\theta_p}{\sigma_p^2}\right)^2}{2\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_p^2}\right)}}, \quad \text{let } A = 2\left(\frac{1}{\sigma} + \frac{1}{\sigma_p}\right)$$

$$\propto e^{\frac{-X_r^2}{2\sigma^2}} e^{\frac{X_r^2}{A\sigma^4}} e^{\frac{X_r\theta_p}{A\sigma^2\sigma_p^2}} \underbrace{e^{\frac{\theta_p^2}{2A\sigma_p^4}}}_{\text{drops out}}$$

$$\propto e^{\overbrace{\frac{\theta p}{A\sigma^2\sigma_p^2}}^{u} X_r - \overbrace{\left(\frac{1}{2\sigma^2} + \frac{1}{A\sigma^4}\right)}^{v} X_r^2}$$

$$A\sigma^2\sigma_p^2 = \sigma^2\sigma_p^2\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_p^2}\right) = \sigma_p^2 + \sigma^2, \quad A\sigma^4 = \sigma^4\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_p^2}\right) = \frac{\sigma_p^4}{\sigma^2} + \sigma_p^2$$

$$\propto N\left(\frac{u}{2v}, \frac{1}{2v}\right) = N\left(\frac{\frac{\theta p}{\sigma_p^2 + \sigma^2}}{}, \right.$$

$$2v = \frac{1}{\sigma^2} + \frac{1}{\frac{\sigma_p^4}{\sigma^2} + \sigma_p^2} = \frac{1}{\sigma^2} + \frac{\sigma^2}{\sigma_p^4 + \sigma^2\sigma_p^2}$$

Scanned with CamScanner