Math 341 lec 15

want to show what information this prior have

$\vec{F}$: iid normal(θ, σ^2) with σ^2 known $P(\theta|\sigma^2) = N(\mu_0, \tau^2) = N(\mu_0, \frac{\sigma^2}{n_0})$ prior represent n_0 pseudodata are distributed $N(\mu_0, \frac{\sigma^2}{n_0})$

Imagine ~~psed~~ pseudodata $Y_1, \dots, Y_{n_0} \overset{iid}{\sim} N(\mu_0, \sigma^2) \Rightarrow \bar{Y} \sim N(\mu_0, \frac{\sigma^2}{n_0})$

$\Rightarrow P(\theta|X, \sigma^2) = N(\hat{\theta}_p, \sigma_p^2)$ where $\hat{\theta}_p = \dfrac{n\bar{x}/\sigma^2 + \mu_0/\tau^2}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}} = \dfrac{n\bar{X} + n_0\mu_0}{n + n_0}$

conjugate

$\frac{\sigma^2}{\tau^2} = n_0$? $\frac{\sigma^2}{n_0} = \tau^2$

$\sigma_p^2 = \dfrac{1}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}} = \dfrac{\sigma^2}{n + n_0}$ (like the variance average) as n go large $\sigma_p^2 \to 0$

$n^* = 1$

convolution?

deriving predictive posterior $P(X^*|X, \sigma^2) = \int_\Theta P(X^*|\theta, \sigma^2) P(\theta|x, \sigma^2)\, d\theta = \int_{\mathbb{R}} N(\theta, \sigma^2) N(\hat{\theta}_p, \sigma_p^2)\, d\theta$

$= \int_{\mathbb{R}} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x^*-\theta)^2}{2\sigma^2}}\right)\left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma_p^2}(\theta - \hat{\theta}_p)^2}\right) d\theta$

$\propto \int_{\mathbb{R}} e^{-\frac{(X^*-\theta)^2}{2\sigma^2}} e^{-\frac{(\theta-\hat{\theta}_p)^2}{2\sigma_p^2}}\, d\theta = \int_{\mathbb{R}} e^{\underbrace{-(\frac{1}{2\sigma^2} + \frac{1}{2\sigma_p^2})}_{-b}\theta^2 + \underbrace{(\frac{x^*}{\sigma^2} + \frac{\hat{\theta}_p}{\sigma_p^2})}_{a}\theta - \frac{x^{*2}}{2\sigma^2} - \frac{\hat{\theta}_p^2}{2\sigma_p^2}}\, d\theta$

$N\left(\frac{x^*\sigma_p^2 + \hat{\theta}_p\sigma^2}{\sigma^2 + \sigma_p^2}, \frac{\sigma^2\sigma_p^2}{\sigma^2 + \sigma_p^2}\right) \sim$ $\propto \int_{\mathbb{R}} e^{\overbrace{(-\frac{1}{2\sigma^2} - \frac{1}{2\sigma_p^2})}^{-b}\theta^2 + \overbrace{(\frac{x^*}{\sigma^2} + \frac{\hat{\theta}_p}{\sigma_p^2})}^{a}\theta}\, d\theta \cdot e^{-\frac{x^{*2}}{2\sigma^2}}$

$= e^{-\frac{x^{*2}}{2\sigma^2}} \int_{\mathbb{R}} e^{-\left(\frac{\sigma_p^2 + \sigma^2}{2\sigma^2\sigma_p^2}\right)\theta^2 + \left(\frac{x^*\sigma_p^2 + \hat{\theta}_p\sigma^2}{\sigma^2\sigma_p^2}\right)\theta}\, d\theta$

$a^2 = \left(\frac{x^*\sigma_p^2 + \hat{\theta}_p\sigma^2}{\sigma^2\sigma_p^2}\right)^2$

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

$\frac{a^2}{4b} = \frac{(x^*\sigma_p^2 + \hat{\theta}_p\sigma^2)^2}{2(\sigma^2 + \sigma_p^2)\sigma^2\sigma_p^2}$

$b = \frac{\sigma^2 + \sigma_p^2}{2\sigma^2\sigma_p^2}$

if $\theta \sim N(\frac{a}{2b}, \frac{1}{2b}) = \dfrac{1}{\sqrt{2\pi(\frac{1}{2b})}} e^{-\frac{1}{2(\frac{1}{2b})}(x - \frac{a}{2b})^2} = \sqrt{\frac{b}{\pi}}\, e^{-b\theta^2 + a\theta - \frac{a^2}{4b}} = \underbrace{\sqrt{\frac{b}{\pi}} e^{-\frac{a^2}{4b}}}_{c} \underbrace{e^{-b\theta^2 + a\theta}}_{k(\theta)}$

$\propto e^{a\theta - b\theta^2}$

$= e^{-\frac{x^{*2}}{2\sigma^2}} \sqrt{\frac{\pi}{b}} e^{\frac{a^2}{4b}} \underbrace{\int \sqrt{\frac{b}{\pi}} e^{-\frac{a^2}{4b}} e^{a\theta - b\theta^2}\, d\theta}_{=1} = e^{-\frac{x^{*2}}{2\sigma^2}} \sqrt{\frac{\pi}{b}}\, e^{\frac{a^2}{4b}}$ → notice that a is a function of x^*, b is not, a constant get rid of it.

put in a^2 ⇒ $a^2 = \left(\frac{x^*\sigma_p^2 + \hat{\theta}_p\sigma^2}{\sigma^2\sigma_p^2}\right)^2$

$\propto e^{-\frac{x^{*2}}{2\sigma^2}} e^{\left(\frac{x^*}{\sigma^2} + \frac{\hat{\theta}_p}{\sigma_p^2}\right)^2/4b} = e^{-\frac{x^{*2}}{2\sigma^2}} e^{\frac{x^{*2}}{4\sigma^4 b}} e^{\frac{x^*\hat{\theta}_p}{2\sigma^2\sigma_p^2 b}} e^{\frac{\hat{\theta}_p^2}{4\sigma_p^4 b}}$

just expand it

$\propto e^{\underbrace{\frac{\hat{\theta}_p}{2\sigma^2\sigma_p^2 b}}_{A} x^* - \underbrace{(\frac{1}{2\sigma^2} - \frac{1}{4\sigma^4 b})}_{B} x^{*2}} = e^{Ax^* - Bx^{*2}} \propto N\left(\frac{A}{2B}, \frac{1}{2B}\right) = N(\hat{\theta}_p, \sigma_p^2 + \sigma^2)$

recording 42min

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

$\frac{1}{2B} = \dfrac{1}{2(\frac{1}{2\sigma^2} - \frac{1}{4\sigma^4 b})} = \dfrac{2\sigma^4 b}{2\sigma^2 b - 1}$ ② put in ① $= \sigma_p^2 + \sigma^2$

$\frac{A}{2B} = A(\sigma_p^2 + \sigma^2) = \dfrac{\hat{\theta}_p}{2b\sigma^2\sigma_p^2}(\sigma_p^2 + \sigma^2) = \dfrac{\hat{\theta}_p}{(1 + \frac{\sigma^2}{\sigma_p^2})\sigma_p^2}(\sigma_p^2 + \sigma^2) = \hat{\theta}_p$

$\therefore \frac{A}{2B} = \hat{\theta}_p$

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

n: # of iid x_i n_0 : # of pseudodata.

law of large # theorem

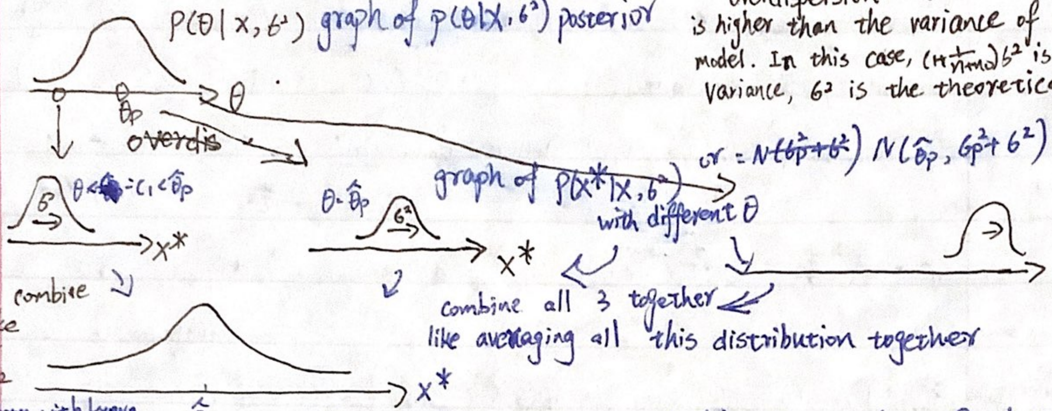
prior $\theta \sim N(\mu_0, \tau^2)$ $\because \sigma_p^2 = \frac{\sigma^2}{n+n_0}$ if $n \to$ large $\hat{\theta}_p \to \bar{x} \to \theta$

remember $P(\theta|X,\sigma^2) = N(\hat{\theta}_p, \sigma_p^2)$ $\hat{\theta}_p \to \theta$, $\sigma_p^2+\sigma^2 \to \sigma^2$

$$P(X^*|X,\sigma^2) = N(\hat{\theta}_p, \sigma_p^2+\sigma^2) = N(\hat{\theta}_p, (1+\frac{1}{n+n_0})\sigma^2) \longrightarrow N(\theta, \sigma^2)$$

overdispersion: when the observed variance is higher than the variance of the theoretical model. In this case, $(1+\frac{1}{n+n_0})\sigma^2$ is observed variance, σ^2 is the theoretical variance.

overdispersion



recording 48min

an overdispersed norm with known variance is normal

higher variance (but never grow around variance) , the $\hat{\sigma}_p^2$ is due to , unknown θ and estimate it. if get more data ie $\hat{\sigma}_p^2 \to 0$ then I'm more sure about θ then we get $\to N(\theta,\sigma^2)$

$\tilde{F}$: iid $N(\theta,\sigma^2)$ θ known   looking for inference of σ^2

likelihood first: $\mathcal{L}(\sigma^2; X, \theta) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\Sigma(x_i-\theta)^2}$

$l(\sigma^2; X,\theta) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\Sigma(x_i-\theta)^2$

regard σ^2 as 1 unit $l'(\sigma^2; X,\theta) = \cancel{-\frac{n}{\sigma^2}+\frac{4}{\sigma^3}[\Sigma(x_i-\theta)^2]}\, -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}(\Sigma(x_i-\theta)^2) = 0$

$\frac{\Sigma(x_i-\theta)^2}{n} = \sigma^2 = \hat{\sigma}^2_{MLE}$ $n\hat{\sigma}^2_{MLE} = \Sigma(x_i-\theta)^2$ (actually is a constant)

$S^2 = \frac{\Sigma(x_i-\bar{x})^2}{n-1} \longrightarrow \sigma^2$

58-60min Laplace prior : $P(\sigma^2|\theta) \propto 1$ (equally) no special σ^2

improper since $\sigma^2 \in (0,+\infty)$ you can't put a constant in infinity

$\lambda e^{-\lambda x}$

$x^k e^{-\frac{kc}{x}}$

$P(\sigma^2|X,\theta) \propto P(X|\theta,\sigma^2) = (2\pi\sigma^2)^{-n/2} e^{-\frac{n\hat{\sigma}^2_{MLE}}{2\sigma^2}} \propto (\sigma^2)^{-\frac{n}{2}} e^{-\frac{n\hat{\sigma}^2_{MLE}/2}{\sigma^2}}$

remember

$U \sim \text{Gamma}(\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} u^{\alpha-1} e^{-\beta u}$ if $v = \frac{1}{u} = t(u) \sim P\cancel{(t^{-1}(v))}\, f_u(t^{-1}(v)) \left|\frac{d}{dv}[t^{-1}(v)]\right|$

$= \frac{\beta^\alpha}{\Gamma(\alpha)}(\frac{1}{v})^{\alpha-1} e^{-\frac{\beta}{v}} \frac{1}{v^2} = \frac{\beta^\alpha}{\Gamma(\alpha)}(\frac{1}{v})^{\alpha+1} e^{-\frac{\beta}{v}}$: Inverse gamma (α,β) abbr InvGamma(α,β)

$\propto v^{-\alpha-1} e^{-\frac{\beta}{v}}$

so $(\sigma^2)^{-\frac{n}{2}} e^{-\frac{\frac{n}{2}\hat{\sigma}^2_{MLE}}{\sigma^2}}$ is InvGamma$(\frac{n}{2}-1, \frac{n}{2}\hat{\sigma}^2_{MLE})$ $\frac{n-2}{2}$

conclusion: under Laplace prior , posterior is inverse Gamma

1:16:20

lec 15-3

conjugate prior: Conjugacy?

$P(\sigma^2|X,\theta) \propto P(X|\theta,\sigma^2)P(\sigma^2|\theta) \propto (\sigma^2)^{-\frac{n}{2}} e^{-\frac{\frac{n}{2}\hat{\sigma}^2_{MLE}}{\sigma^2}} \quad k(\sigma^2|\theta) \quad k(\sigma^2|\theta) \overset{?}{=} ?$

$= (\sigma^2)^a e^{-\frac{b}{\sigma^2}}$ ~~invgamma(-a-1, b)~~ is the conjugate prior.

$= (\sigma^2)^{-\frac{n}{2}+a} e^{-\frac{\frac{n}{2}\hat{\sigma}^2_{MLE}+b}{\sigma^2}}$

let $P(\sigma^2|\theta) = \text{InvGamma}(\alpha, \beta) \Rightarrow P(\sigma^2|x,\theta) \propto \left((\sigma^2)^{-\frac{n}{2}} e^{-\frac{\frac{n}{2}\hat{\sigma}^2_{MLE}}{\sigma^2}}\right)\left((\sigma^2)^{-\alpha-1} e^{-\frac{\beta}{\sigma^2}}\right)$

$= \cancel{f(\sigma^2)} (\sigma^2)^{-(\frac{n}{2}+\alpha)-1} e^{-\frac{(\frac{n}{2}\hat{\sigma}^2_{MLE}+\beta)}{\sigma^2}} \quad \propto \text{InvGamma}(\frac{n}{2}+\alpha, \frac{n}{2}\hat{\sigma}^2_{MLE}+\beta)$

want another parameterization to ~~be easily~~ interpretate easily via pseudo count.

let $\alpha = \frac{n_0}{2}$ remember $\hat{\sigma}^2_{MLE} = \Sigma(x_i-\theta)^2$, $\beta = \frac{n_0}{2}\sigma_0^2 \Rightarrow P(\sigma^2|\theta) = \text{InvGamma}(\frac{n_0}{2}, \frac{n_0\sigma_0^2}{2})$

$\Rightarrow P(\sigma^2|\theta, x) = \text{InvGamma}(\frac{n+n_0}{2}, \frac{n\hat{\sigma}^2_{MLE}+n_0\sigma_0^2}{2})$

pseudodata: $Y_1, \dots, Y_{n_0} \sim N(\underset{\text{known}}{\theta}, \underset{\text{belief}}{\sigma_0^2}) \Rightarrow n_0\sigma_0^2 = \Sigma(Y_i-\theta)^2 \quad \sigma_0^2 = \frac{\Sigma(Y_i-\theta)^2}{n_0}$

imagine you see n_0 data point if n_0 small, $\rightarrow$ it's uninformative.

Haldane : $n_0 = 0$, $\sigma_0^2 = ?$ $\Rightarrow P(\sigma^2|\theta) = \text{InvGamma}(0,0)$

will go through laplace Jeffery