Math 341 4/2/19 Lec 16

$$p(x^*|x,\sigma^2) \propto e^{-\frac{x^{*2}}{2\sigma^2}} \sqrt{\frac{\pi}{b}}\, e^{\frac{a^2}{4b}} \propto e^{-\frac{x^{*2}}{2\sigma^2} + \frac{a^2}{4b}} = e^{-\frac{x^{*2}}{2\sigma^2} + \frac{x^{*2}}{4b(\sigma^2)^2} + \frac{2x^*\theta_p}{4b\sigma^2\sigma_p^2}}\, e^{\frac{\theta_p^2}{(\sigma_p^2)^2 4b}}$$

constant wrt x^*

$$a^2 = \left(\frac{x^*}{\sigma^2} + \frac{\theta_p}{\sigma_p^2}\right)^2 = \frac{x^{*2}}{(\sigma^2)^2} + \frac{2x^*\theta_p}{\sigma^2\sigma_p^2} + \frac{\theta_p^2}{(\sigma_p^2)^2}, \quad b = \frac{1}{2\sigma^2} + \frac{1}{2\sigma_p^2}$$

$$\cancel{\propto e^{\underbrace{\left(\frac{2\theta_p}{4b\sigma^2\sigma_p^2} - \frac{1}{2\sigma^2}\right)}_{c} x^* - \underbrace{\left(-\frac{1}{4b\sigma^2\sigma_p^2}\right)}_{d} x^{*2}} \propto N\left(\frac{c}{2d}, \frac{1}{2d}\right)}$$

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

$$\cancel{\frac{1}{2d} = \frac{1}{2\left(-\frac{1}{4b\sigma^2\sigma_p^2}\right)} = -2b\sigma^2\sigma_p^2 = -2\left(\frac{1}{2\sigma^2} + \frac{1}{2\sigma_p^2}\right)\sigma^2\sigma_p^2 = \sigma_p^2 + \sigma^2}$$

$$\cancel{\frac{c}{2d} = c\left(\frac{1}{2d}\right) = \left(\frac{\theta_p}{2b\sigma^2\sigma_p^2} - \frac{1}{2\sigma^2}\right)(\sigma_p^2 + \sigma^2) = \left(\frac{\theta_p}{\sigma_p^2 + \sigma^2} - \frac{1}{2\sigma^2}\right)(\sigma_p^2 + \sigma^2) = \theta_p - \frac{\sigma_p^2 + \sigma^2}{2\sigma^2}}$$

$$\cancel{4b\sigma^2\sigma_p^2 = 4\left(\frac{1}{2\sigma^2} + \frac{1}{2\sigma_p^2}\right)\sigma^2\sigma_p^2 = \sigma_p^2 + \sigma^2}$$

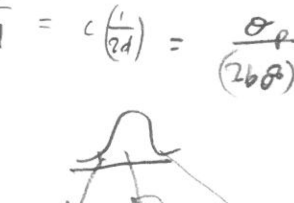
$$\cancel{4b(\sigma^2)^2 = 4\left(\frac{1}{2\sigma^2} + \frac{1}{2\sigma_p^2}\right)(\sigma^2)^2 = 2\left(\sigma^2 + \frac{(\sigma^2)^2}{\sigma_p^2}\right)}$$

$$\propto e^{\underbrace{\frac{\theta_p}{2b\sigma^2\sigma_p^2}}_{c} x^* - \underbrace{\left(\frac{1}{2\sigma^2} - \frac{1}{4b(\sigma^2)^2}\right)}_{d} x^{*2}} \propto N\left(\frac{c}{2d}, \frac{1}{2d}\right) = N(\theta_p, \sigma^2 + \sigma_p^2)\ !!$$

$$\frac{1}{2d} = \frac{1}{2\left(\frac{1}{2\sigma^2} - \frac{1}{4b(\sigma^2)^2}\right)} = \frac{1}{\frac{1}{\sigma^2} - \frac{1}{2b(\sigma^2)^2}} = \frac{1}{\frac{2b(\sigma^2)^2 - \sigma^2}{2b(\sigma^2)^2}} = \frac{2b(\sigma^2)^2}{2b(\sigma^2)^2 - \sigma^2} = \frac{2b(\sigma^2)^2}{2b\sigma^2 - 1} = \frac{2\left(\frac{1}{2\sigma^2} + \frac{1}{2\sigma_p^2}\right)(\sigma^2)^2}{2\left(\frac{1}{2\sigma^2} + \frac{1}{2\sigma_p^2}\right)\sigma^2 - 1}$$

$$= \frac{\sigma^2 + \frac{(\sigma^2)^2}{\sigma_p^2}}{1 + \frac{\sigma^2}{\sigma_p^2} - 1} = \frac{\sigma^2\sigma_p^2 + (\sigma^2)^2}{\sigma^2} = \sigma_p^2 + \sigma^2$$

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



$$\sigma_p^2 \xrightarrow{n\to\infty} 0 \qquad p(x^*|x,\sigma^2) \xrightarrow{n\to\infty} p(x^*|\theta,\sigma^2)$$

(2

We are done with inferring θ when σ^2 known. Now we do the opposite: inferring σ^2 when θ known.

$F: \underbrace{X_1, \ldots, X_n}_{X} \mid \theta, \sigma^2 \stackrel{iid}{\sim} N(\theta, \sigma^2)$

$$P(X \mid \theta, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(X_i - \theta)^2} = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\Sigma(X_i-\theta)^2}$$

$= \mathcal{L}(\sigma^2; X, \theta)$

Let's find $\hat{\sigma}^2_{MLE}$ first.

$$\ell(\sigma^2; X, \theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\underbrace{\Sigma(X_i-\theta)^2}_{\text{constant!}}$$

$$\ell'(\sigma^2; X, \theta) = -\frac{n}{\cancel{2}}\frac{1}{\sigma^2} + \frac{1}{\cancel{2}(\sigma^2)^2}\Sigma(X_i-\theta)^2 \stackrel{set}{=} 0$$

$$\Rightarrow -n + \frac{\Sigma(X_i-\theta)^2}{\sigma^2} = 0 \Rightarrow \hat{\sigma}^2_{MLE} = \frac{1}{n}\Sigma(X_i-\theta)^2$$

Make sense?
Avg. sqd deviation from mean!!

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

a kernel we haven't seen!

Back to prob. theory... 621 stuff...

Let's say $P(Y) \propto Y^{-a} e^{-\frac{b}{Y}}$

Let us try to solve for the normalization constant

$$\frac{1}{c} = \int_0^\infty u^{-a} e^{-b/u}\, du$$

let $t = \frac{b}{u} \Rightarrow u = \frac{b}{t} \Rightarrow \frac{du}{dt} = -bt^{-2} \Rightarrow du = -bt^{-2}dt$

$u=0 \Rightarrow t=\infty, \quad u=\infty \Rightarrow t=0$

$$\Rightarrow \frac{1}{c} = \int_{\infty}^{0} \left(\frac{b}{t}\right)^{-a} e^{-t}\left(-bt^{-2}\right)dt = \int_0^{\infty} b^{-a} t^{a} e^{-t} b t^{-2} dt$$

$$= \int_0^{\infty} b^{-a+1} t^{a-2} e^{-t} dt = b^{-a+1}\int_0^{\infty} t^{(a-1)-1} e^{-t} dt = b^{-a+1}\Gamma(a-1)$$

$$\Rightarrow c = \frac{b^{a-1}}{\Gamma(a-1)}$$

$\Uparrow$ InvGamma conjugate prior for likelihood model

$$\Rightarrow P(y) = \frac{b^{a-1}}{\Gamma(a-1)} y^{-a} e^{-\frac{b}{y}} = \text{InvGamma}(a+1, b)$$

$\alpha \leq 1 \Rightarrow$ no mean!!

In general ... $P(Y) = \text{InvGamma}(\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-\alpha-1} e^{-\frac{\beta}{y}}$

$\text{Supp}(Y) = (0, \infty)$, Param space: $\alpha, \beta > 0$

$E[Y] = \frac{\beta}{\alpha-1}$, $\text{Mode}(Y) = \frac{\beta}{\alpha+1}$, $\text{Med}(Y) = \text{qinvgamma}(0.5, \alpha, \ldots$

why called inverse gamma?

$P(U) = \text{Gamma}(\alpha, \beta)$

$V = t(U) = \frac{1}{U} \sim ?$

$\Updownarrow$

$u = t^{-1}(v) = \frac{1}{v}$

$$f_V(v) = f_U(t^{-1}(v))\left|\frac{d}{dv}\left[t^{-1}(v)\right]\right| = f_U\left(\frac{1}{v}\right)\left|\frac{d}{dv}\left[\frac{1}{v}\right]\right|$$

$$= \left(\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{v}\right)^{\alpha-1} e^{-\beta\frac{1}{v}}\right)\left|-\frac{1}{v^2}\right| = \frac{\beta^\alpha}{\Gamma(\alpha)} v^{-\alpha+1-2} e^{-\frac{\beta}{v}}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} v^{-\alpha-1} e^{-\frac{\beta}{v}}$$

$\mathcal{F}: X_1, \ldots, X_n \mid \theta, \sigma^2 \overset{iid}{\sim} N(\theta, \sigma^2)$ with θ known

$P(\sigma^2 \mid \theta) = \text{InvGamma}(\alpha, \beta)$

$$\Rightarrow P(\sigma^2 \mid X, \theta) \propto \underbrace{\left((\sigma^2)^{-n/2} e^{-\frac{n\hat{\sigma}^2_{mle}/2}{\sigma^2}}\right)}_{k(X\mid\theta)} \underbrace{\left((\sigma^2)^{-\alpha-1} e^{-\frac{\beta}{\sigma^2}}\right)}_{k(\sigma^2\mid\theta)} = \underbrace{(\sigma^2)^{-\left(\frac{n}{2}+\alpha\right)-1} e^{-\frac{\frac{n\hat{\sigma}^2_{mle}}{2}+\beta}{\sigma^2}}}_{k(\sigma^2\mid X,\theta)}$$

$$\propto \text{InvGamma}\left(\frac{n}{2}+\alpha,\ \frac{n\hat{\sigma}^2_{MLE}}{2}+\beta\right)$$

units don't really line up so no easy pseudocount interpretation.

Let's parametrize the prior differently.

$$P(\sigma^2) = \text{InvGamma}\left(\frac{n_0}{2}, \frac{n_0\sigma_0^2}{2}\right) \propto (\sigma^2)^{-\frac{n_0}{2}-1} e^{-\frac{n_0\sigma_0^2/2}{\sigma^2}}$$

prior exp: $E(\sigma^2) = \frac{\frac{n_0\sigma_0^2}{2}}{\frac{n_0}{2}-1} = \frac{n_0}{n_0-2}\sigma_0^2$

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

Pseudocount Interpretation:

n_0 : # pseudobservations (strength)

σ_0^2 : variance of the pseudobservations

Pt. Est.

$$\Rightarrow \hat{\sigma}^2_{MMSE} = \frac{n\hat{\sigma}^2_{MLE}+n_0\sigma_0^2}{n+n_0-2}$$

$$\hat{\sigma}^2_{MAP} = \frac{n\hat{\sigma}^2_{MLE}+n_0\sigma_0^2}{n+n_0+2}$$

$$\hat{\sigma}^2_{MMAE} = \text{QInvGamma}\left(0.5, \frac{n+n_0}{2}, \frac{n\hat{\sigma}^2_{MLE}+n_0\sigma_0^2}{2}\right)$$

CR's & Hypothesis Tests... same!!

Shrinkage

$$\hat{\sigma}^2_{MMSE} = \frac{n\hat{\sigma}^2_{MLE}}{n+n_0-2} + \frac{n_0\sigma_0^2}{n+n_0-2}\cdot\frac{n_0-2}{n_0-2}$$

$$= \underbrace{\frac{n}{n+n_0-2}}_{1-c}\hat{\sigma}^2_{MLE} + \underbrace{\frac{n_0-2}{n+n_0-2}}_{c}E(\sigma^2)$$

<u>Uninformative Priors</u>

① Laplace $P(\sigma^2 \mid \theta) \propto 1$

$$P(\sigma^2 \mid \theta, x) \propto P(X \mid \sigma^2, \theta) \propto (\sigma^2)^{-\left(\frac{n}{2}+1\right)+1} e^{-\frac{n\hat{\sigma}^2_{MLE}/2}{\sigma^2}} \propto \text{InvGamma}\left(\frac{n-2}{2}, \frac{n\hat{\sigma}^2_{MLE}}{2}\right)$$

Only proper if $n \geq 3$

$\Rightarrow P(\sigma^2 \mid \theta) = \text{InvGamma}(-1, 0)$ i.e. improper