Lec 21 Math 341/641

$X_1, \dots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$, σ^2 known

$$f(\theta|\sigma^2) = N(\mu_0, \tau^2) \Rightarrow f(\theta|X, \sigma^2) = N\Bigg(\underbrace{\frac{\frac{\mu_0}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}}_{\theta_p}, \underbrace{\frac{1}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}}_{\sigma_p^2}\Bigg)$$

$$\hat{\theta}^{mmse} = \hat{\theta}^{mmae} = \hat{\theta}^{map} = \theta_p$$

Pseudodata interpretation of hyperparams?

Imagine pseudodata $y_1, \dots, y_{n_0}$

$$\text{let } \mu_0 = \bar{y} = \sum_{i=1}^{n_0} y_i, \text{ let } \tau^2 = \frac{\sigma^2}{n_0} \Rightarrow f(\theta|\sigma^2) = N\left(\bar{y}, \frac{\sigma^2}{n_0}\right)$$

Two Hyperparams: $n_0, \bar{y}$

What does posterior look like?

$$\sigma_p^2 = \frac{1}{\frac{1}{\tau^2} + \frac{1}{\sigma^2}} = \frac{1}{\frac{n_0}{\sigma^2} + \frac{n}{\sigma^2}} = \frac{1}{\frac{n_0+n}{\sigma^2}} = \frac{\sigma^2}{n_0+n}$$ Makes sense!

$$\theta_p = \frac{\frac{\mu_0}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}} = \frac{\frac{\bar{y}}{\cancel{\frac{\sigma^2}{n_0}}} + \frac{\bar{x}}{\cancel{\frac{\sigma^2}{n}}}}{\cancel{\frac{n_0+n}{\sigma^2}}} = \frac{n_0\bar{y} + n\bar{x}}{n_0+n} = \frac{\sum y_i + \sum x_i}{n_0+n}$$ overall avg of real and pseudodata

Laplace prior: $f(\theta|\sigma^2) \propto 1 \Rightarrow f(\theta|\sigma^2) = N(0, \infty) \Rightarrow f(\theta|X, \sigma^2) = N\left(\bar{x}, \frac{\sigma^2}{n}\right)$

Haldane prior: $n_0 = 0, \bar{y} = 0 \Rightarrow f(\theta|\sigma^2) = N(0, \infty) \Rightarrow f(\theta|X, \sigma^2) = N\left(\bar{x}, \frac{\sigma^2}{n}\right)$

Jeffrey's Prior...

$$\mathcal{L}(\theta; \vec{x}, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\theta)^2} = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum(x_i-\theta)^2} = (2\pi\sigma^2)^{-n/2} e^{-\frac{\sum x_i^2}{2\sigma^2}} e^{\frac{n\bar{x}\theta}{\sigma^2}} e^{-\frac{\theta^2}{2\sigma^2}}$$

$$\ell(\theta; \vec{x}, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{\sum x_i^2}{2\sigma^2} + \frac{n\bar{x}}{\sigma^2}\theta - \frac{1}{2\sigma^2}\theta^2$$

$$\ell'(\theta; \vec{x}, \sigma^2) = \frac{n\bar{x}}{\sigma^2} - \frac{1}{\sigma^2}\theta$$

$$\ell''(\theta; \vec{x}, \sigma^2) = -\frac{1}{\sigma^2}$$

$$I_n(\theta|\sigma^2) = \mathbb{E}[-\ell''] = \frac{1}{\sigma^2}$$

$$f_J(\theta|\sigma^2) \propto \sqrt{I_n(\theta)} = \sqrt{\frac{1}{\sigma^2}} \propto 1 \Rightarrow f_J(\theta|\sigma^2) = N(0, \infty) \Rightarrow f(\theta|X, \sigma^2)\ N(\bar{x}, \cdot$$

Shrinkage

$$\hat{\theta}^{MMSE} = \frac{\Sigma y_i + \Sigma x_i}{n_0 + n} = \frac{\Sigma y_i}{n_0+n} \cdot \frac{n_0}{n_0} + \frac{\Sigma x_i}{n_0+n} \cdot \frac{n}{n} = \underbrace{\frac{n_0}{n_0+n}}_{\rho} \underbrace{\bar{y}}_{E(\theta)} + \underbrace{\frac{n}{n_0+n}}_{(1-\rho)} \underbrace{\bar{x}}_{\hat{\theta}^{MLE}}$$

Posterior Predictive Distr. let $n_* = 1$

MATH 340 Lec 19

$$f(X_* \mid \vec{X}, \sigma^2) = \int_{\mathbb{R}} \underbrace{f(X_* \mid \theta, \sigma^2)}_{N(\theta, \sigma^2)} \underbrace{f(\theta \mid \vec{X}, \sigma^2)}_{N(\theta_p, \sigma_p^2)} d\theta = N(\theta_p, \sigma_p^2 + \sigma^2)$$

For $n_* \geq 1 \Rightarrow$ MVN!

$X_1, \ldots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$ θ known, σ^2 param space is $(0, \infty)$

$$f(\sigma^2 \mid \vec{X}, \theta) = \frac{f(\vec{X} \mid \theta, \sigma^2) f(\sigma^2 \mid \theta)}{f(\vec{X} \mid \theta)}$$

Pseudodata interpretation?
n_0: # of pseudopoints, $y_1, \ldots, y_{n_0}$
$\sigma_0^2 = \frac{1}{n_0} \sum_{i=1}^{n_0} (y_i - \theta)^2$
MLE for σ^2 on pseudodata

$$\propto f(\vec{X} \mid \theta, \sigma^2) f(\sigma^2 \mid \theta)$$

$$= \left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i - \theta)^2} \right) f(\sigma^2 \mid \theta)$$

$$= (2\pi)^{-n/2} (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\Sigma(x_i-\theta)^2} f(\sigma^2 \mid \theta)$$

Pattern match

$$\propto (\sigma^2)^{-n/2} e^{-\frac{\Sigma(x_i-\theta)^2/2}{\sigma^2}} (\sigma^2)^{-\alpha-1} e^{-\frac{\beta}{\sigma^2}} \Rightarrow f(\sigma^2 \mid \theta) = \text{InvGamma}(\alpha, \beta)$$

$$= (\sigma^2)^{-(\alpha + \frac{n}{2}) - 1} e^{-\frac{\beta + \Sigma(x_i-\theta)^2/2}{\sigma^2}}$$

$$\propto \text{InvGamma}\left(\alpha + \frac{n}{2}, \beta + \frac{\Sigma(x_i-\theta)^2}{2}\right)$$

let $\alpha = \frac{n_0}{2}$, $\beta = \frac{n_0 \sigma_0^2}{2}$

$$\Rightarrow f(\sigma^2 \mid \vec{X}, \sigma^2) = \text{InvGamma}\left(\frac{n_0+n}{2}, \frac{n_0 \sigma_0^2 + \Sigma(x_i-\theta)^2}{2}\right) \Rightarrow f(\sigma^2 \mid \theta) = \text{InvGamma}\left(\frac{n_0}{2}, \frac{n_0\sigma_0^2}{2}\right)$$

Review of Inverse Gamma

$$Y \sim \text{InvGamma}(\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-\alpha-1} e^{-\frac{\beta}{y}} \mathbb{1}_{y\in(0,\infty)}$$

$$E(Y) = \int_0^\infty y \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-\alpha-1} e^{-\frac{\beta}{y}} dy = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty y^{-\alpha} e^{-\frac{\beta}{y}} dy$$

let $u = \frac{1}{y} \Rightarrow y = \frac{1}{u}$, $\frac{dy}{du} = -\frac{1}{u^2} \Rightarrow dy = -\frac{1}{u^2} du$

$y = 0 \Rightarrow u = \infty$, $y = \infty \Rightarrow u = 0$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_\infty^0 u^\alpha e^{-\beta u}\left(-\frac{1}{u^2}\right) du = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty u^{\alpha-1-1} e^{-\beta u} du$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha-1)}{\beta^{\alpha-1}} = \frac{\cancel{\beta^\alpha}\,\cancel{\Gamma(\alpha-1)}}{(\alpha-1)\cancel{\Gamma(\alpha-1)}\,\cancel{\beta^\alpha}\beta^{-1}} = \frac{\beta}{\alpha-1}$$

$$\text{Mode}[Y] = \text{argmax}\left\{\frac{\beta^\alpha}{\Gamma(\alpha)} y^{-\alpha-1} e^{-\frac{\beta}{y}}\right\} = \text{argmax}\left\{y^{-\alpha-1} e^{-\frac{\beta}{y}}\right\}$$

$$= \text{argmax}\left\{(-\alpha-1)\ln(y) - \frac{\beta}{y}\right\}$$

$$\frac{d}{dy}[\quad] = \frac{-(\alpha+1)}{y} + \frac{\beta}{y^2} \overset{set}{=} 0 \Rightarrow \frac{\alpha+1}{\cancel{y}} = \frac{\beta}{y^{\cancel{2}}} \Rightarrow y = \frac{\beta}{\alpha+1}$$ always true

MLE?

$$\hat{\sigma}^{2\,MMSE} = \frac{\frac{n_0\sigma_0^2 + \sum(x_i-\theta)^2}{2}}{\frac{n_0+n}{2}-1} = \frac{n_0\sigma_0^2 + \sum(x_i-\theta)^2}{n_0+n-2}$$

$$\hat{\sigma}^{2\,MMAE} = \text{qinvgamma}\left(.5, \frac{n_0+n}{2}, \frac{n_0\sigma_0^2 + \sum(x_i-\theta)^2}{2}\right)$$

$$\hat{\sigma}^{2\,MAP} = \frac{\frac{n_0\sigma_0^2 - \sum(x_i-\theta)^2}{2}}{\frac{n_0+n}{2}+1} = \frac{n_0\sigma_0^2 + \sum(x_i-\theta)^2}{n_0+n+2}$$

Shrinkage?

from HW but we will prove again now

$$\hat{\sigma}^{2\,MMSE} = \frac{n_0\sigma_0^2 + \sum(x_i-\theta)^2}{n_0+n-2} = \frac{n_0\sigma_0^2}{n_0+n-2}\cdot\frac{n_0-2}{n_0-2} + \frac{\sum(x_i-\theta)^2}{n_0+n-2}\cdot\frac{n}{n} = \underbrace{\frac{n_0-2}{n_0+n-2}}_{\rho}\underbrace{\frac{n_0\sigma_0^2}{n_0-2}}_{E(\sigma^2|\theta)} + \underbrace{\frac{n}{n_0+n-2}}_{1-\rho}\underbrace{\frac{\sum(x_i-\theta)^2}{n}}_{\hat{\sigma}^{2\,MLE}}$$

Laplace Prior

let $f(\sigma^2|\theta) \propto 1$

If we do this...

$$f(\sigma^2|\vec{x},\theta) \propto (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum(x_i-\theta)^2}$$

$$= (\sigma^2)^{-(n/2-1)-1} e^{-\frac{1}{2\sigma^2}\sum(x_i-\theta)^2}$$

$$\propto \text{InvGamma}\left(\frac{n}{2}-1, \frac{\sum(x_i-\theta)^2}{2}\right)$$

$$= \text{InvGamma}\left(\frac{n-2}{2}, \frac{\sum(x_i-\theta)^2}{2}\right) \Rightarrow n_0 = -2, \sigma_0^2 = 0$$

$$\Rightarrow f(\sigma^2|\theta) = \text{InvGamma}(-1, 0) \text{ improper!}$$

strange!

Not used! Flat is weird

under Laplace

$$\hat{\sigma}^{2,MAP} = \hat{\sigma}^{2,MLE} = \frac{n_0\sigma_0^2 + \sum(x_i-\theta)^2}{n_0+n-2} = \frac{\sum(x_i-\theta)^2}{n}$$

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

$$\Rightarrow f(\sigma^2|\theta,\vec{x}) = \text{InvGamma}\left(\frac{n}{2}, \frac{\sum(x_i-\theta)^2}{2}\right)$$ popular

Jeffreys Prior

$$\ell(\sigma^2;\vec{x},\theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{\sum(x_i-\theta)^2}{2\sigma^2}$$

$$\ell'(\sigma^2;\vec{x},\theta) = -\frac{n/2}{\sigma^2} + \frac{\sum(x_i-\theta)^2/2}{(\sigma^2)^2}$$

$$\ell''(\sigma^2;\vec{x},\theta) = \frac{n/2}{(\sigma^2)^2} - \frac{\sum(x_i-\theta)^2}{(\sigma^2)^3}$$

sum of iid χ^2 expectation = 1

$$I(\sigma^2|\theta) = E[-\ell''] = E\left[-\frac{n/2}{(\sigma^2)^2} + \frac{\sum\left(\frac{x_i-\theta}{\sigma}\right)^2}{(\sigma^2)^2}\right] = -\frac{n/2}{(\sigma^2)^2} + \frac{1}{(\sigma^2)^2}E\left[\sum\left(\frac{x_i-\theta}{\sigma}\right)^2\right]$$

$$= \frac{1}{(\sigma^2)^2}\left(-\frac{n}{2}+n\right) = \frac{n}{(\sigma^2)^2}\left(-\frac{1}{2}+1\right) = \frac{n}{2}(\sigma^2)^{-2}$$

$$f_J(\sigma^2|\theta) \propto \sqrt{I(\sigma^2|\theta)} = \sqrt{\frac{n}{2}(\sigma^2)^{-2}} \propto (\sigma^2)^{-1} \propto \text{InvGamma}(0,0)$$ i.e. same as Haldane

$$CR_{\sigma^2, 1-\alpha} = \Big[\text{qinvgamma}\Big(\frac{\alpha}{2}, \frac{n_0+n}{2}, \frac{n_0\sigma_0^2 + \Sigma(x_i-\theta)^2}{2}\Big),$$
$$\text{qinvgamma}\Big(\frac{\alpha}{2}, \frac{n_0+n}{2}, \frac{n_0\sigma_0^2 + \Sigma(x_i-\theta)^2}{2}\Big) \Big]$$

Hypothesis Tests:

$H_a: \sigma^2 > s \Rightarrow H_0: \sigma^2 \le s$

$$\text{pval} = \text{pinvgamma}\Big(s, \frac{n_0+n}{2}, \frac{n_0\sigma_0^2 + \Sigma(x_i-\theta)^2}{2}\Big)$$

$H_a: \sigma^2 < s \Rightarrow H_0: \sigma^2 \ge s$

$$\text{pval} = 1 - \text{pinvgamma}\Big(s, \frac{n_0+n}{2}, \frac{n_0\sigma_0^2 + \Sigma(x_i-\theta)^2}{2}\Big)$$

MID 2 1(g)

Fact from MATH 340

$$X \sim T_k \propto \left(1+\frac{x^2}{k}\right)^{-\frac{k+1}{2}}, \quad Y = a+bX, \quad a\in\mathbb{R}, b>0$$

$$Y \sim f_Y(y) = \frac{1}{b} f_X\left(\frac{x-a}{b}\right) \propto k_X\left(\frac{x-a}{b}\right) = \left(1+\frac{\left(\frac{x-a}{b}\right)^2}{k}\right)^{-\frac{k+1}{2}} = \left(1+\frac{(x-a)^2}{kb^2}\right)^{-\frac{k+1}{2}} \propto T_k(a, b^2)$$

kernel of general T dist with k df.

Posterior Prediction Distribution for $n_* = 1$

$$a = \frac{n_0+n}{2}, \quad b = \frac{n_0\sigma_0^2 + \sum(x_i-\theta)^2}{2}$$

$$p(x_* \mid \vec{x}, \theta) = \int_{\mathbb{R}} \underbrace{p(x_* \mid \theta, \sigma^2)}_{N(\theta,\sigma^2)} \underbrace{p(\sigma^2 \mid \vec{x}, \theta)}_{\text{InvGamma}(a,b)} \, d\sigma^2$$

$$= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_*-\theta)^2} \frac{b^a}{\Gamma(a)} (\sigma^2)^{-a-1} e^{-\frac{b}{\sigma^2}} d\sigma^2$$

$$\propto \int_{\mathbb{R}} (\sigma^2)^{-a-\frac{1}{2}-1} e^{-\frac{1}{2\sigma^2}(x_*-\theta)^2 - \frac{b}{\sigma^2}} d\sigma^2$$

$$Y \sim \text{InvGamma}(c,d) = \frac{d^c}{\Gamma(c)} y^{-c-1} e^{-\frac{d}{y}}$$

$$\Rightarrow \int_0^\infty y^{-c-1} e^{-\frac{d}{y}} dy = \frac{\Gamma(c)}{d^c}$$

$$= \int_{\mathbb{R}} (\sigma^2)^{-(a+\frac{1}{2})-1} e^{-\frac{\left((x_*-\theta)^2/2 + b\right)}{\sigma^2}} d\sigma^2$$

$$= \frac{\Gamma(a-\frac{1}{2})}{\left(\frac{(x_*-\theta)^2}{2} + b\right)^{a+\frac{1}{2}}}$$

$$\propto \left(\frac{(x_*-\theta)^2 + 2b}{2}\right)^{-(a+\frac{1}{2})} \cdot \left(\frac{2}{2b}\right)^{-(a+\frac{1}{2})}$$

$$= \left(1+\frac{(x_*-\theta)^2}{2b}\right)^{-(a+\frac{1}{2})} = \left(1+\frac{(x_*-\theta)^2}{2b}\right)^{-\frac{\overbrace{n_0+n}^{k}+1}{2}}$$

Student's T is another type of an overdispersed normal

$$= \left(1+\frac{1}{n_0+n}\frac{(x_*-\theta)^2}{\frac{2b}{n_0+n}}\right)^{-\frac{n_0+n+1}{2}} \propto T_{n_0+n}\left(\theta, \frac{2b}{n_0+n}\right) = T_{n+n_0}\left(\theta, \frac{n_0\sigma_0^2 + \sum(x_i-\theta)^2}{n_0+n}\right)$$