Lec 18 4/19

n_0 = ghost observations

σ_0^2 = best guess at σ^2 in ghost obs.

<u>Jeffrey's Prior</u>: $P_J(\sigma^2|\theta) \propto \sqrt{I(\sigma^2;\theta)}$

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

$$I(\sigma^2,\theta) = E_X[-\ell''] = E_X\left[\frac{-n}{2(\sigma^2)^2} + \frac{\sum(x_i-\theta)^2}{(\sigma^2)^3}\right] = \left(\frac{-n}{2(\sigma^2)^2} + \frac{\sum E[(x_i-\theta)^2]}{(\sigma^2)^3}\right)$$

$$\boxed{\sum E[(x_i-\theta)^2] = VAR[X]}$$

$X \sim$ since normal dist $\rightarrow \sigma^2$

$$\boxed{\sum\sigma^2 = n\sigma^2}$$

$$= \left(\frac{-n}{2(\sigma^2)^2} + \frac{\sum\sigma^2}{(\sigma^2)^3}\right) = \boxed{\frac{n}{2(\sigma^2)^2}}$$

$$P_J(\sigma^2|\theta) \propto \sqrt{\frac{n}{2(\sigma^2)^2}} \propto \sqrt{\frac{1}{(\sigma^2)^2}} = (\sigma^2)^{-1} \propto \text{InvGamma}(0,0)$$

Jeffrey = InvGamma(0, 0) = (Same as Haldane)

These (Haldane / Jeffrey) are the default prior for this model.

<u>Shrinkage</u>: generally the prior is by $P(\sigma^2|\theta) = \text{InvGamma}\left(\frac{n_0}{2}, \frac{n_0\sigma_0^2}{2}\right)$

$$\Rightarrow E[\sigma^2|\theta] = \frac{n_0\sigma_0^2}{n_0-2} \quad (n_0 > 2) \text{ we must}$$

$$\hat{\sigma}^2_{MMSE} = \frac{n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}{n+n_0-2} = \underbrace{\frac{n\hat{\sigma}^2_{MLE}}{n+n_0-2}}_{1-Q} + \frac{n_0\sigma_0^2}{n+n_0-2}\cdot\frac{n_0-2}{n_0-2}$$

n_0 large $\Rightarrow$ shrink hard to prior as if we have seen a lot of ghost observations

$$= \underbrace{\frac{n}{n+n_0-2}}_{1-Q}\hat{\sigma}^2_{MLE} + \underbrace{\frac{n_0-2}{n+n_0-2}}_{Q} E[\sigma^2|\theta]$$

Posterior Predictive Distribution:

cancel anything which is not a function of X_*, σ^2

$$P(X_* | X, \theta) = \int_0^\infty \underbrace{P(X_* | \theta, \sigma^2)}_{N(\theta, \sigma^2)} \underbrace{P(\sigma^2 | X, \theta)}_{\text{InvGamma}\left(\frac{n+n_0}{2}, \frac{n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}{2}\right)} d\sigma^2$$

with $\alpha = \frac{n+n_0}{2}$ and $\beta = \frac{n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}{2}$

$$= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(X_*-\theta)^2} \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} e^{-\frac{\beta}{\sigma^2}}$$

$$\propto \int_0^\infty (\sigma^2)^{-\frac{1}{2}} e^{-\frac{(X_*-\theta)^2/2}{\sigma^2}} (\sigma^2)^{-\alpha-1} e^{-\frac{\beta}{\sigma^2}} d\sigma^2$$

$$= \int_0^\infty (\sigma^2)^{-\overbrace{(\alpha+\frac{1}{2})}^{\alpha'}-1} e^{-\frac{\overbrace{(X_*-\theta)^2/2+\beta}^{\beta'}}{\sigma^2}} d\sigma^2$$

$$\boxed{\underline{\text{u-sub}} \quad \frac{\Gamma(\alpha')}{\beta'^{\alpha'}} = \Gamma(\alpha')\beta'^{-\alpha'}}$$

$$\propto \Gamma\left(\frac{n+n_0+1}{2}\right) \boxed{\left(\frac{(X_*-\theta)^2 + n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}{2}\right)^{-\frac{n+n_0+1}{2}}}$$

$$P(X_* | X, \theta) \propto$$

$$\boxed{\begin{aligned} \text{Let } v &= n+n_0 \\ a &= n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2 \end{aligned}}$$

$$\propto \left(\frac{(X_*-\theta)^2 + a}{2}\right)^{-\frac{v+1}{2}} \cdot \underbrace{\left(\frac{2}{a}\right)^{-\frac{v+1}{2}} \cdot \left(\frac{2}{a}\right)^{\frac{v+1}{2}}}_{\text{same game 7/7}}$$

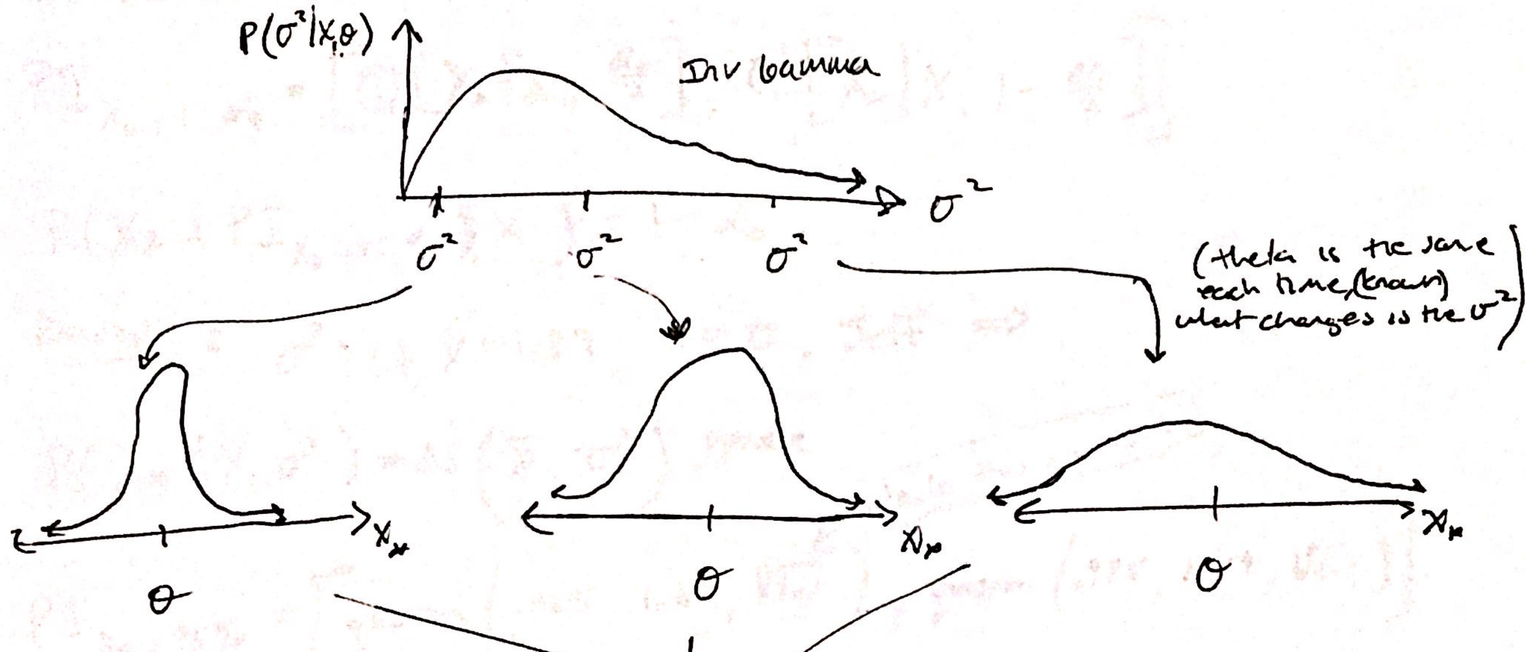
$$= \left(\frac{(X_*-\theta)^2}{a} + 1\right)^{-\frac{v+1}{2}} \left(\frac{2}{a}\right)^{\frac{v+1}{2}}$$

$$\left(\frac{\frac{1}{v}}{\frac{1}{v}}\right)$$

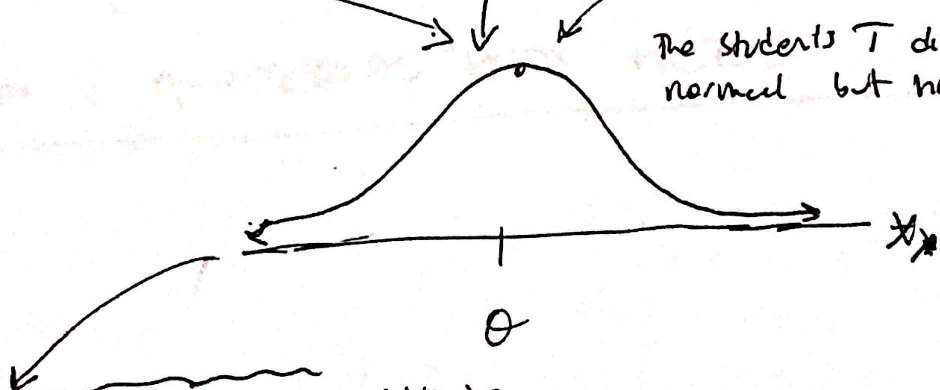
$$= \left(1+\frac{1}{v}\frac{(x_*-\theta)^2}{\frac{a}{v}}\right)^{-\frac{v+1}{2}}\left(\frac{2}{a}\right)^{\frac{v+1}{2}} \propto \left(1+\frac{1}{v}\frac{(x_*-\theta)^2}{\frac{a}{v}}\right)^{-\frac{v+1}{2}}$$

$$\propto T_v\left(\theta, \frac{a}{v}\right) = T_{n+n_0}\left(\theta, \frac{n\hat{\sigma}^2_{MLE}+n_0\sigma_0^2}{n+n_0}\right)$$

This dist is the "non-standard Student's T distribution" or "shifted and scaled Student's T distribution".



The students T dist looks like a normal but has thicker tails.



(n large $\to \hat{\sigma}^2_{MLE} \to \sigma^2$)

$$T_{n+n_0}\left(\theta, \frac{n\hat{\sigma}^2_{MLE}+n_0\sigma_0^2}{n+n_0}\right) \overset{n+n_0>20}{\approx} N\Bigg(\theta, \underbrace{\frac{n\hat{\sigma}^2_{MLE}+n_0\sigma_0^2}{n+n_0}}_{\text{Scale parameter}}\Bigg) \approx N(\theta, \sigma^2)$$

Lec 18 4/19

Some examples: $\theta = 5$, $n = 12$, $\hat{\sigma}^2_{MLE} = .387$, Jeff prior $\Rightarrow n_0 = \sigma_0^2 = 0$

$$P(X_* > 8 \mid X, \theta) = 1 - P(X_* \leq 8 \mid X, \theta)$$

$$= 1 - \underbrace{\text{pt.scaled}}_{\text{for } t \text{ dist}}\left(\underset{X^*}{8},\ \underset{\nu = 12}{n + n_0},\ \underset{5}{\theta},\ \underset{\sqrt{\text{scale parameter}}}{\sqrt{\frac{12 \cdot (.387)}{12}}}\right)$$

$\sqrt{(\cdot)}$

Predictive Intervals (PI)

$$PI_{X_*, 1-\alpha_0} = \left[Q\left[X_* \mid X, \tfrac{\alpha_0}{2}\right], Q\left[X_* \mid X, 1 - \tfrac{\alpha_0}{2}\right]\right]$$

$$P(X_* \in PI_{X_*, 1-\alpha_0} \mid X) = 1 - \alpha_0$$

Example: $\sigma^2 = 1.1$, $\bar{x} = 1.89$, $n = 13$, Jeff $\Rightarrow$

$P(X_* \mid X, \sigma^2) = N(\bar{x}, \sigma^2)$ hence

note this on c.s.

$$PI_{X_*, 95\%} = \left[\text{qnorm}(.025, 1.89, \sqrt{1.1}), \text{qnorm}(.975, 1.89, \sqrt{1.1})\right]$$

MIDTERM 2 MATERIAL ENDS HERE

$\mathcal{F}: X_1, \ldots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$ where both θ, σ^2 are unknown.

thus we want inference for both, or inference for one and the other is a "nuisance parameter". Usual case: interested in mean, drop the one we don't care ~~bab~~ about.

lets assume Laplace prior

$$P(\theta, \sigma^2 | X) \propto P(X|\theta,\sigma^2)P(\theta,\sigma^2) \propto P(X|\theta,\sigma^2)$$

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

$$\propto (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2} \; ?$$

$\propto$ represent this another way

$$\sum(X_i-\theta)^2 = \sum((X_i-\bar{X}) + (\bar{X}-\theta))^2 = \sum(X_i-\bar{X})^2 + 2\sum(X_i-\bar{X})(\bar{X}-\theta) + \sum(\bar{X}-\theta)^2$$

$\boxed{\text{Sample variance } s^2 := \frac{1}{n-1}\sum(X_i-\bar{X})^2} \Rightarrow = (n-1)s^2 + 2\sum(X_i\bar{X} - \bar{X}^2 - X_i\theta + \bar{X}\theta) + n(\bar{X}-\theta)^2$

$\bar{X} = \frac{1}{n}\sum X_i$

$$= (n-1)s^2 + n(\bar{X}-\theta)^2 + 2(\cancel{n\bar{X}^2} - \cancel{n\bar{X}^2} \cancel{+ n\bar{X}\theta} + \cancel{n\bar{X}\theta})$$

$$= (n-1)s^2 + n(\bar{X}-\theta)^2 \quad \text{So.}$$

$$P(\theta, \sigma^2|X) \propto (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\left((n-1)s^2 + n(\bar{X}-\theta)^2\right)}$$

$$= (\sigma^2)^{-(\frac{n}{2}+1)-1} e^{-\frac{(n-1)s^2/2}{\sigma^2}} e^{-\frac{1}{2\sigma^2} n(\bar{X}-\theta)^2}$$

$$= \underbrace{e^{-\frac{1}{2\frac{\sigma^2}{n}}(\bar{X}-\theta)^2}}_{\propto N(\bar{X}, \frac{\sigma^2}{n})} \underbrace{(\sigma^2)^{-(\frac{n}{2}+1)-1} e^{\frac{-(n-1)s^2/2}{\sigma^2}}}_{\propto \text{InvGamma}\left(\frac{n+2}{2}, \frac{(n-1)s^2}{2}\right)}$$

$$\propto \text{Normal Inverse Gamma}\left(\bar{X}, \frac{\sigma^2}{n}, \frac{n+2}{2}, \frac{(n-1)s^2}{2}\right) \leftarrow (\mu, \lambda, \alpha, \beta)$$