Jeffrey's $\quad P_J(\sigma^2 \mid \theta) \propto \sqrt{I(\sigma^2; \theta)}$

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

$$\ell''(\sigma^2; X, \theta) = \frac{n}{2(\sigma^2)^2} - \frac{\Sigma(X_i-\theta)^2}{(\sigma^2)^3}$$

$X \sim N(\theta, \sigma^2)$

Def of $Var[X] = \sigma^2$

$$I(\sigma^2; \theta) = E_X[-\ell''] = E_X\left[-\frac{n}{2(\sigma^2)^2} + \frac{\Sigma(X_i-\theta)^2}{(\sigma^2)^3}\right]$$

Fisher Info

$$= \left(-\frac{n}{2(\sigma^2)^2} + \frac{\Sigma E[(X_i-\theta)^2]}{(\sigma^2)^3}\right) = \left(-\frac{n}{2(\sigma^2)^2} + \frac{\Sigma\sigma^2}{(\sigma^2)^3}\right)$$

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

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

Jeffrey's and Haldane are equivalent principled objective priors and they're the default for this model.

Shrinkage $\quad P(\sigma^2 \mid \theta) = \text{InvGamma}\left(\frac{n_0}{2}, \frac{n_0\sigma_0^2}{2}\right) \Rightarrow E[\sigma^2 \mid \theta] = \frac{n_0\sigma_0^2}{n_0-2}$

(Valid for $n_0 > 2$)

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

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

Posterior Predictive Distribution

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

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

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

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

$$\overset{u\text{-subst}}{=} \frac{\Gamma(\alpha')}{\beta'^{\alpha'}} = \Gamma(\alpha')\beta'^{-\alpha'}$$

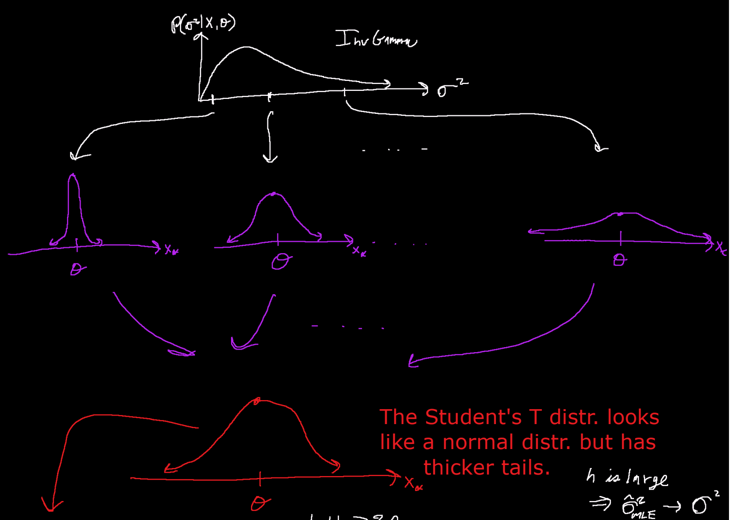
$$P(X_* \mid X, \theta) \propto \Gamma\left(\frac{n+n_0+1}{2}\right)\left(\frac{(X_*-\theta)^2 + \overbrace{n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}^{q}}{2}\right)^{-\frac{\overbrace{n+n_0}^{\nu}+1}{2}}$$

$$\propto \left(\frac{(X_*-\theta)^2+q}{2}\right)^{-\frac{\nu+1}{2}} \cdot \left(\frac{2}{q}\right)^{-\frac{\nu+1}{2}} \cdot \left(\frac{2}{q}\right)^{\frac{\nu+1}{2}}$$

$$= \left(\frac{(X_*-\theta)^2}{q} \cdot \frac{\frac{1}{\nu}}{\frac{1}{\nu}} + 1\right)^{-\frac{\nu+1}{2}} \left(\frac{2}{q}\right)^{\frac{\nu+1}{2}} = \left(1 + \frac{1}{\nu}\frac{(X_*-\theta)^2}{\frac{q}{\nu}}\right)^{-\frac{\nu+1}{2}} \left(\frac{2}{q}\right)^{\frac{\nu+1}{2}}$$

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

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



n is large $\Rightarrow \hat{\sigma}^2_{MLE} \to \sigma^2$

$n+n_0 > 20$

$$T_{n+n_0}\left(\theta, \frac{n\hat{\sigma}^2_{MLE}+n_0\sigma_0^2}{n+n_0}\right) \approx N\left(\theta, \frac{n\hat{\sigma}^2_{MLE}+n_0\sigma_0^2}{n+n_0}\right) \approx N(\theta, \sigma^2)$$

$\theta = 5$, $n = 12$, $\hat{\sigma}^2_{MLE} = 0.387$, Jeffrey's prior $\Rightarrow n_0 = \sigma_0^2 = 0$

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

$$= 1 - \text{pt.scaled}\Big(\underbrace{8}_{X_*}, \underbrace{12}_{\nu}, \underbrace{5}_{\theta}, \underbrace{\sqrt{\tfrac{\cancel{12}\cdot 0.387}{\cancel{12}}}}_{\sqrt{\text{scale}}}\Big)$$

Predictive Intervals (PI)

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

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

$\sigma^2 = 1.1$, $\bar{x} = 1.89$, $n = 13$, Jeffrey's prior $\Rightarrow P(X_* \mid X, \sigma^2) = N(\bar{x}, \sigma^2)$

$$PI_{X_*, 95\%} = [\text{qnorm}(.025, 1.89, \sqrt{1.1}), \text{qnorm}(0.975, 1.89, \sqrt{1.1})]$$

MID 2 ⇑

FINAL ⇓

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

thus we want inference for both or inference for one and the other is a "nuisance parameter"

Let's assume Laplace prior

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

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

$$\propto (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\Sigma(X_i-\theta)^2} \overset{?}{\propto} \text{InvGamma}$$

It is not invgamma since the posterior now is a 2-d distr and the invgamma is only one-dimensional

What we have above is a known distribution but to get it into canonical form, we need to do some algebra

$$\Sigma(X_i-\theta)^2 = \Sigma((X_i-\bar{x})+(\bar{x}-\theta))^2 = \Sigma(X_i-\bar{x})^2 + 2\Sigma(X_i-\bar{x})(\bar{x}-\theta) + \Sigma(\bar{x}-\theta)^2$$

$s^2 := \frac{1}{n-1}\Sigma(x_i-\bar{x})^2$ the sample variance formula from Math 241

$\bar{x} = \frac{1}{n}\Sigma x_i \Rightarrow \Sigma x_i = n\bar{x}$

$$= (n-1)s^2 + 2\Sigma(x_i\bar{x} - \bar{x}^2 - x_i\theta + \bar{x}\theta) + n(\bar{x}-\theta)^2$$

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

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

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

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

$$\propto \text{NormalInvGamma}\left(\mu = \bar{x}, \lambda = n, \alpha = \frac{n+2}{2}, \beta = \frac{(n-1)s^2}{2}\right)$$

This is the "normal-inverse-gamma" distribution with four parameters!