2D Normal with Theta / Sigma unknown.

Integrate out the nuisance parameter. Now longest derivation:

Assume Jeffrey's Prior $P(\theta, \sigma^2) \propto \frac{1}{\sigma^2}$

$$P(X_*|X) \propto \int_0^\infty \int_{\mathbb{R}} k(X_*|\theta,\sigma^2)\, k(\theta,\sigma^2|X)\, d\theta\, d\sigma^2$$

$$= \int_0^\infty \int_{\mathbb{R}} \frac{1}{\sqrt{\sigma^2}} e^{-\frac{1}{2\sigma^2}(X_*-\theta)^2} (\sigma^2)^{-\frac{n}{2}-1} e^{-\frac{(n-1)s^2/2}{\sigma^2}} e^{-\frac{n}{2\sigma^2}(\theta-\bar{x})^2}\, d\theta\, d\sigma^2$$

Re-organize the terms, not a function of θ gets taken out.

$$= \int_0^\infty (\sigma^2)^{-\frac{1}{2}} (\sigma^2)^{-\frac{n}{2}-1} e^{-\frac{(n-1)s^2/2}{\sigma^2}} \int_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}\underbrace{\left((X_*-\theta)^2 + n(\theta-\bar{x})^2\right)}_{X_*^2-(2X_*\theta)+(n+1)\theta^2 -2n\theta\bar{x}+n\bar{x}^2}}\, d\theta\, d\sigma^2$$

$$= \int_0^\infty (\sigma^2)^{-\frac{n+1}{2}-1} \underbrace{e^{-\frac{(n-1)s^2/2 + X_*^2/2 + n\bar{x}^2/2}{\sigma^2}}} \int_{\mathbb{R}} \underbrace{e^{+\frac{X_*+n\bar{x}}{\sigma^2}\theta - \frac{n+1}{2\sigma^2}\theta^2}}_{e^{a\theta-b\theta^2}}\, d\theta\, d\sigma^2$$

$$\int_{\mathbb{R}} e^{a\theta-b\theta^2}\, d\theta = \sqrt{\frac{\pi}{b}}\, e^{a^2/4b}$$

$$= \int_0^\infty (\sigma^2)^{-\frac{n+1}{2}-1} \boxed{} \sqrt{\frac{\pi}{\frac{n+1}{2\sigma^2}}}\, e^{\underbrace{\left(\frac{X_*+n\bar{x}}{\sigma^2}\right)^2 / 4\left(\frac{n+1}{2\sigma^2}\right)}_{\frac{(X_*+n\bar{x})^2}{2\sigma^2(n+1)}}}$$

$$= \int_0^\infty (\sigma^2)^{-\frac{n+1}{2}-1} e^{-\frac{\overbrace{(n-1)s^2/2 + x_*^2/2 + n\bar{x}^2/2 - (x_* + n\bar{x})^2/(2(n+1))}^{\beta}}{\sigma^2}} (\sigma^2)^{1/2}\sqrt{\frac{2\pi}{n+1}}\, d\sigma^2$$

(trying to derive posterior predictive distribution)

$$\propto \int_0^\infty (\sigma^2)^{-\overbrace{\frac{n}{2}}^{\alpha}-1} e^{-\beta/\sigma^2}\, d\sigma^2 = \Gamma(\alpha)\beta^{-\alpha} = \Gamma\left(\tfrac{n}{2}\right)\beta^{-\alpha} \propto \beta^{-\alpha}$$

$$= \left(\frac{(n-1)s^2}{2} + \frac{x_*^2}{2} + \frac{n\bar{x}^2}{2} - \frac{\overbrace{(x_* + n\bar{x})^2}^{x_*^2 + 2nx_*\bar{x} + n^2\bar{x}^2}}{2n+2}\right)^{-n/2}$$

$$= (ax_*^2 + bx_* + c)^{-n/2}$$

$$a = \frac{1}{2} - \frac{1}{2n+2} = \frac{1}{2}\left(1 - \frac{1}{n+1}\right) = \frac{1}{2}\,\frac{n}{n+1}$$

$$b = -\frac{2n\bar{x}}{2n+2} = -\frac{n\bar{x}}{n+1}$$

$$c = \frac{(n-1)s^2}{2} + \frac{n\bar{x}^2}{2} - \frac{n^2\bar{x}^2}{2n+2} = \frac{1}{2}\left(\frac{(n-1)s^2 + n\bar{x}^2 - n^2\bar{x}^2}{n+1}\right.$$

$$= \cancel{a}\left(\frac{1}{a}\right)^{n/2}\left(\frac{1}{a}\right)^{-n/2}(ax_*^2 + bx_* + c)^{-n/2} = \left(\frac{1}{a}\right)^{-n/2}\left(x_*^2 - \frac{b}{a}x_* + \frac{c}{a}\right)^{-n/2}$$

$$\propto \left(x_*^2 + \frac{b}{a}x_* + \frac{c}{a}\right)^{-n/2}$$ (reverse complete the square)

$$\propto \left(\left(x_* + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + \frac{c}{a}\right)^{-n/2} \cdot \left(\frac{1}{\frac{c}{a} - \frac{b^2}{4a^2}}\right)^{-n/2}$$

since prop multiply by anything.

$$\propto \left(1 + \frac{\left(x_* + \frac{b}{2a}\right)^2}{\frac{c}{a} - \frac{b^2}{4a^2}}\right)^{-\frac{\overbrace{(n-1)}^{\nu}+1}{2}} = \left(1 + \frac{1}{n-1}\,\frac{\left(x^* - \overbrace{\left(-\frac{b}{2a}\right)}^{\mu}\right)^2}{\underbrace{\left(\frac{c}{a} - \frac{b^2}{4a^2}\right)(n-1)}_{S_0^2}}\right)^{-\frac{(n-1)+1}{2}}$$

$$\boxed{-\frac{b}{2a} = \frac{\frac{n\bar{x}}{n+1}}{\frac{n}{n+1}} = \bar{X}} = \mu$$

$$\& \; T_v(\mu, s_0^2) = T_{n-1}\left(\bar{X}, \left(\frac{n+1}{n}\right)s^2\right) \underset{n \text{ large}}{\approx} N(\bar{X}, s^2)$$

$$\frac{c}{a} = \frac{\frac{1}{2}\left((n-1)s^2 + n\bar{x}^2 - \frac{n^2\bar{x}^2}{n+1}\right)}{\frac{1}{2}\,\frac{n}{n+1}}$$

$$\frac{c}{a} = \frac{(n-1)(n+1)}{n}s^2 + (n+1)\bar{x}^2 - n\bar{x}^2$$

$$-\frac{b^2}{4a^2} = -\left(\frac{b}{2a}\right)^2 = -\left(-\frac{b}{2a}\right)^2 = -\bar{x}^2$$

$$\frac{c}{a} - \frac{b^2}{4a^2} = \frac{(n-1)(n+1)}{n}s^2$$

$$s_0^2 = \left(\frac{n+1}{n}\right)s^2$$

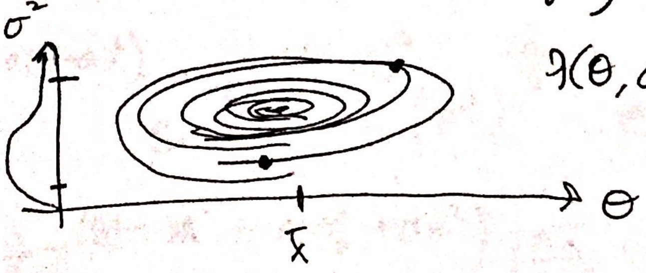
def of conditional probability.

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

Under Jeff prior

$$\text{Norm Inv Gamma} = N\left(\bar{x}, \frac{\sigma^2}{n}\right)\ \text{InvGamma}\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$$

You can think of a normal inv gamma as first sampling from an inverse gamma$\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$ to get a σ^2 value, and then that σ^2 is used to draw a θ from $N(\bar{x}, \frac{\sigma^2}{n})$ and return the vector $\begin{bmatrix}\theta \\ \sigma^2\end{bmatrix}$.



$g(\theta, \sigma^2 | X)$

This can also be done another way....

$$P(\theta,\sigma^2|X) = \underbrace{P(\sigma^2|X,\theta)}_{\text{InvGamma}\left(\frac{n}{2}, \frac{n\hat{\sigma}^2_{MLE}}{2}\right)} \underbrace{P(\theta|X)}_{T_{n-1}\left(\bar{x}, s^2/2\right)}$$

If we decompose the first way, we draw theta from $N(\bar{x}, \frac{\sigma^2}{n})$ So σ^2 must be known. What if we break this lay, instead of Self prior, use $N(\mu_0, \tau^2) = P(\theta)$? and $P(\sigma^2) = \text{InvGamma}\left(\frac{n_0}{2}, \frac{n_0\sigma_0^2}{2}\right)$ These were the two priors we began with when we started investigating the normal likelihood model. However, its important to note we are not allowing $\tau^2 = \sigma^2/n_0$. What happens?

$P(\theta,\sigma^2) = P(\theta)P(\sigma^2)$ not $P(\theta|\sigma^2)P(\sigma^2)$ The two priors are disconnected.

Lets derive the posterior under this 2D prior.

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

$$\propto k(x|\theta,\sigma^2)\,k(\theta)\,k(\sigma^2) =$$

$$= (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}((n-1)s^2 + n(\bar{x}-\theta)^2)} e^{-\frac{1}{2\tau^2}(\theta-\mu_0)^2} (\sigma^2)^{-\frac{n_0}{2}-1} e^{-\frac{n_0\sigma_0^2/2}{\sigma^2}}$$

$$= (\sigma^2)^{-\frac{n}{2}-\frac{n_0}{2}-1} e^{-\frac{1}{2\sigma^2}((n-1)s^2 + n_0\sigma_0^2 + n\bar{x}^2)} e^{\left(\frac{n\bar{x}}{\sigma^2}+\frac{\mu_0}{\tau^2}\right)\theta - \left(\frac{n}{2\sigma^2}+\frac{1}{2\tau^2}\right)\theta^2}$$

pick up next class, here.