More insight into conjugate prior for (μ, σ^2)?

- Suppose μ unknown, & σ^2 known $y|\mu,\sigma^2 \sim N(\mu,\sigma^2)$

- PRIOR - $p(\mu) = N(\mu_0, \tau_0^2)$ $\qquad n > 1$

$$p(\mu|y) = N\left(\frac{\frac{1}{\tau_0^2}}{\frac{1}{\tau_0^2}+\frac{n}{\sigma^2}} \mu_0 + \frac{\frac{n}{\sigma^2}}{\frac{1}{\tau_0^2}+\frac{n}{\sigma^2}} \bar{y}, \; \frac{1}{\frac{1}{\tau_0^2}+\frac{n}{\sigma^2}} \right)$$

We could have decided to let τ_0^2 be a function of the known σ^2

$$\tau_0^2 = \frac{\sigma^2}{\kappa_0}$$ so that we set κ_0 instead of τ_0^2

$\frac{1}{\sigma^2/\kappa}$

then:

$$p(\mu|y) = N\left(\frac{\frac{\kappa_0}{\sigma^2}}{\frac{\kappa_0}{\sigma^2}+\frac{n}{\sigma^2}} \mu_0 + \frac{\frac{n}{\sigma^2}}{\frac{\kappa_0}{\sigma^2}+\frac{n}{\sigma^2}} \bar{y}, \; \frac{1}{\frac{\kappa_0}{\sigma^2}+\frac{n}{\sigma^2}} \right)$$

So, now everything is in terms of σ^2 & we can think about κ_0 analogously to n. This provides a way in some situations to more easily think about amount of prior information we want to include.

We can simplify the above $p(\mu|y)$ as

$$= N\left(\frac{\kappa_0}{\kappa_0+n} \mu_0 + \frac{n}{\kappa_0+n} \bar{y}, \; \frac{\sigma^2}{\kappa_0+n} \right)$$

(HOPE THIS HELPS!)