2020.2.11

$$p(\theta,\sigma^2|y) \propto p(y|\theta,\sigma^2)\, p(\theta,\sigma^2)$$

$$\propto p(y|\theta,\sigma^2)\, \underline{p(\theta|\sigma^2)}\, \underline{p(\sigma^2)}$$

$$\propto (\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{(n-1)s^2}{2\sigma^2}\right] \exp\left[-\frac{n(\theta-\bar{y})^2}{2\sigma^2}\right] \leftarrow \text{likelihood}$$

$$\cdot \frac{1}{\sqrt{2\pi\frac{\sigma^2}{n_0}}} \exp\left[-\frac{(\theta-\mu_0)^2}{2\sigma^2/n_0}\right] \leftarrow \theta \quad \text{prior of } \theta|\sigma^2$$

$$\cdot \frac{\left(\frac{\nu_0\sigma_0^2}{2}\right)^{\frac{\nu_0}{2}}}{\Gamma\left(\frac{\nu_0}{2}\right)} (\sigma^2)^{-\frac{\nu_0}{2}-1} \exp\left[-\frac{\frac{\nu_0\sigma_0^2}{2}}{\sigma^2}\right] \leftarrow \text{prior of } \sigma^2$$

Step 1: Make a whole square for θ inside the exponent:

$$\frac{n(\theta-\bar{y})^2}{2\sigma^2} + \frac{n_0(\theta-\mu_0)^2}{2\sigma^2}$$

$$= \frac{1}{2\sigma^2}\left[(n+n_0)\theta^2 - 2(n\bar{y}+n_0\mu_0)\theta + n\bar{y}^2 + n_0\mu_0^2\right]$$

$$= \frac{1}{2\sigma^2}\left\{(n+n_0)\left[\theta - \frac{n\bar{y}+n_0\mu_0}{n+n_0}\right]^2 - \frac{(n\bar{y}+n_0\mu_0)^2}{n+n_0} + n\bar{y}^2 + n_0\mu_0^2\right\}$$

$$= \frac{1}{2\sigma^2}\left\{(n+n_0)\left[\theta - \frac{n\bar{y}+n_0\mu_0}{n+n_0}\right]^2 + \frac{-n^2\bar{y}^2 - n_0^2\mu_0^2 - 2nn_0\bar{y}\mu_0 + n^2\bar{y}^2 + nn_0\bar{y}^2 + nn_0\mu_0^2 + n_0^2\mu_0^2}{n+n_0}\right.$$

$$= \frac{1}{2\sigma^2}\left\{(n+n_0)\left[\theta - \frac{n\bar{y}+n_0\mu_0}{n+n_0}\right]^2 + \frac{nn_0(\bar{y}-\mu_0)^2}{n+n_0}\right\} ✓$$

Step 2: Plug in back to $p(\theta,\sigma^2|y)$ and simplify:

$$p(\theta,\sigma^2|y) \propto (\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{(n-1)s^2}{2\sigma^2}\right]$$

$$\times (\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{n+n_0}{2\sigma^2}\left(\theta - \frac{n\bar{y}+n_0\mu_0}{n+n_0}\right)^2\right] \leftarrow \text{a normal density for } \theta$$

$$\times \exp\left[-\frac{1}{2\sigma^2}\cdot\frac{nn_0(\bar{y}-\mu_0)^2}{n+n_0} ✓\right] \leftarrow \text{from Step 1}$$

$$\times (\sigma^2)^{-\frac{\nu_0}{2}-1} \exp\left[-\frac{\nu_0\sigma_0^2}{2\sigma^2}\right]$$

$$\Rightarrow p(\theta,\sigma^2|y) \propto \underbrace{(\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{(n+n_0)}{2\sigma^2}\left(\theta-\frac{n\bar{y}+n_0\mu_0}{n+n_0}\right)^2\right]}_{\text{a conditional density of } \theta \text{ given } \sigma^2}$$

$$\times \underbrace{(\sigma^2)^{-\frac{\nu_0+n}{2}-1} \exp\left\{-\frac{1}{2\sigma^2}\left[(n-1)s^2+\nu_0\sigma_0^2+\frac{nn_0(\bar{y}-\mu_0)^2}{n+n_0}\right]\right\}}_{\text{a marginal density of } \sigma^2}$$

In other words,

$$p(\theta,\sigma^2|y) = p(\theta|\sigma^2,y) \times p(\sigma^2|y)$$

Key idea: If $p(\theta,\sigma^2|y) \propto \tilde{p}(\theta|\sigma^2,y) \times \tilde{p}(\sigma^2|y)$ such that $\tilde{p}(\theta|\sigma^2,y)$ is the "core part" of a density of θ, and $\tilde{p}(\sigma^2|y)$ is the "core part" of a density of σ^2, then we can identify the distributions of

$$\theta|\sigma^2,y \quad \text{and} \quad \sigma^2|y$$

from $\tilde{p}(\theta|\sigma^2,y)$ and $\tilde{p}(\sigma^2|y)$, respectively.

We can conclude that

(1) $$\theta|\sigma^2,y \sim N\left(\frac{n\bar{y}+n_0\mu_0}{n+n_0},\ \frac{\sigma^2}{n+n_0}\right)$$

Since an inverse-gamma density looks like

$$(\sigma^2)^{-\boxed{\cdots}-1} \exp\left(-\frac{\boxed{\cdots}}{\sigma^2}\right)$$

we can identify that

(2) $$\sigma^2|y \sim \text{Inv Gamma}\left(\frac{\nu_0+n}{2},\ \frac{(n-1)s^2+\nu_0\sigma_0^2+\frac{nn_0(\bar{y}-\mu_0)^2}{n+n_0}}{2}\right)$$

(1) + (2): normal-inverse-gamma distribution for (θ,σ^2)

Interpretation:

$$\theta \mid \sigma^2, y \sim N\left(\frac{n\bar{y}+n_0\mu_0}{n+n_0}, \frac{\sigma^2}{n+n_0}\right)$$

$$\text{mean} = \frac{n}{n+n_0}\cdot\bar{y} + \frac{n_0}{n+n_0}\mu_0$$ weighted average

$$\text{Var} = \frac{\sigma^2}{n+n_0},$$ $n+n_0$: sample size + "prior sample size"

$$\sigma^2 \mid y \sim \text{InvGamma}\left(\frac{\nu_0+n}{2}, \frac{1}{2}\left[\nu_0\sigma_0^2 + (n-1)s^2 + \frac{n n_0(\bar{y}-\mu_0)^2}{n+n_0}\right]\right)$$

check the slides for interpretation

- Marginal distribution of $\theta \mid y$

To get the marginal, we integrate out the other variables.

$$p(\theta \mid y) = \int_0^\infty p(\theta, \sigma^2 \mid y)\, d\sigma^2$$

$$= \int_0^\infty p(\theta \mid \sigma^2, y)\, p(\sigma^2 \mid y)\, d\sigma^2$$

Using the notation in the slides,

$$\theta \mid \sigma^2, y \sim N\left(\mu_1, \frac{\sigma^2}{n_1}\right), \quad \sigma^2 \mid y \sim \text{InvGamma}\left(\frac{\nu_1}{2}, \frac{\nu_1\sigma_1^2}{2}\right)$$

$$\Rightarrow p(\theta \mid y) \propto \int_0^\infty \underbrace{(\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{(\theta-\mu_1)^2}{2\sigma^2/n_1}\right)}_{\text{normal } \theta \mid \sigma^2, y}$$

InvGamma $\sigma^2 \mid y$

$$\times (\sigma^2)^{-\frac{\nu_1}{2}-1} \exp\left(-\frac{\nu_1\sigma_1^2}{2\sigma^2}\right) d\sigma^2$$

$$= \int_0^\infty (\sigma^2)^{-\frac{\nu_1+1}{2}-1} \exp\left[-\frac{n_1(\theta-\mu_1)^2 + \nu_1\sigma_1^2}{2\sigma^2}\right] d\sigma^2$$

Change of variable:

$$u = \frac{n_1(\theta-\mu_1)^2 + \nu_1\sigma_1^2}{2\sigma^2}, \quad \sigma^2 = \frac{n_1(\theta-\mu_1)^2 + \nu_1\sigma_1^2}{2u}$$

$$d\sigma^2 = -\frac{n_1(\theta-\mu_1)^2 + \nu_1\sigma_1^2}{2u^2}\, du$$

replace σ^2 with u

$$\int_0^{+\infty} \cdots d\sigma^2 = \int_{+\infty}^0 \cdots \times \frac{d\sigma^2}{du}\cdot du$$

$$p(\theta|y) \propto \int_0^{\infty} (\sigma^2)^{-\frac{\nu_1+1}{2}-1} \exp\left\{-\frac{n_1(\theta-\mu_1)^2+\nu_1\sigma_1^2}{2\sigma^2}\right\} d\sigma^2$$

$$\propto \int_0^{\infty} \left[\frac{2u}{n_1(\theta-\mu_1)^2+\nu_1\sigma_1^2}\right]^{\frac{\nu_1+1}{2}+1} \exp(-u)$$

$$\times \frac{n_1(\theta-\mu_1)^2+\nu_1\sigma_1^2}{2u^2}\, du \qquad \left(-\frac{d\sigma^2}{du}\right)$$

$$\propto \underbrace{\int_0^{\infty} u^{\frac{\nu_1+1}{2}-1} e^{-u}\, du}_{=\Gamma\left(\frac{\nu_1+1}{2}\right) = \text{const (of } \theta)} \times \left[\frac{\nu_1\sigma_1^2+n_1(\theta-\mu_1)^2}{2}\right]^{-\frac{\nu_1+1}{2}}$$

$$\propto \left[\frac{\nu_1\sigma_1^2+n_1(\theta-\mu_1)^2}{2}\right]^{-\frac{\nu_1+1}{2}}$$

$$\propto \left[\nu_1\sigma_1^2+n_1(\theta-\mu_1)^2\right]^{-\frac{\nu_1+1}{2}}$$

$$\propto \left[1+\frac{n_1(\theta-\mu_1)^2}{\nu_1\sigma_1^2}\right]^{-\frac{\nu_1+1}{2}}$$

$$= \left[1+\frac{1}{\nu_1}\left(\frac{\theta-\mu_1}{\sigma_1/\sqrt{n_1}}\right)^2\right]^{-\frac{\nu_1+1}{2}}$$

$$\Rightarrow p(\theta|y) \propto \left[1+\frac{1}{\nu_1}\underbrace{\left(\frac{\theta-\mu_1}{\sigma_1/\sqrt{n_1}}\right)}_{W}\right]^{-\frac{\nu_1+1}{2}}$$

$$\Rightarrow \frac{\theta-\mu_1}{\sigma_1/\sqrt{n_1}} \,\Big|\, y \sim t_{\nu_1}$$

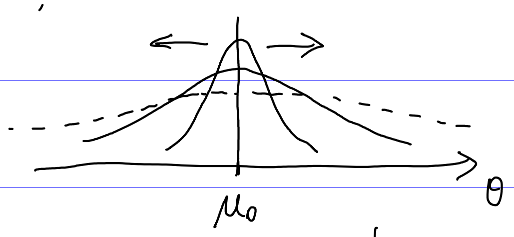
the recentered and rescaled θ has a t distribution.

Improper priors:

A normal prior for θ:

$$p(\theta \mid \mu_0, \tau_0^2) \simeq \frac{1}{\sqrt{2\pi\tau_0^2}} \exp\left\{ -\frac{(\theta-\mu_0)^2}{2\tau_0^2} \right\}$$

If $\tau_0^2 \to +\infty$,



mathematically, $\tau_0^2 \to \infty$, then $\frac{1}{\sqrt{2\pi\tau_0^2}} \to 0$,

$$\exp\left(-\frac{(\theta-\mu_0)^2}{2\tau_0^2}\right) \approx \exp(-0) = 1$$

So $p(\theta) \to 0$

In the limit, $p(\theta)$ is no longer a proper density.

But the posterior is still proper.

- Reference prior: many definitions

One definition: given an assumed model and the observed data, the reference prior is the least informative prior in a certain information-theoretic sense.

Implication: the reference prior can depend on the data!