4 | 27 | 17 HW tonight

$X_1, \dots, X_n \mid \theta, \sigma^2 \overset{iid}{\sim} N(\theta, \sigma^2)$

$P(\theta, \sigma^2) \propto \frac{1}{\sigma^2}$

$\overset{\shortparallel}{P(\theta)P(\sigma^2)}$ are Jeffrey's priors.

You don't know θ or σ^2

$\Rightarrow P(\theta, \sigma^2 \mid X) = \text{Norm Inv Gamma}$

(Before Midterm) $P(\theta \mid X, \sigma^2) = N\left(\bar{x}, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$

You know θ or σ^2 $P(\sigma^2 \mid X, \theta) = \text{Inv Gamma}\left(\frac{n}{2}, \frac{n\hat{\sigma}^2}{2}\right)$

$P(\theta \mid X) = T_{n-1}\left(\bar{x}, \frac{s}{\sqrt{n}}\right)$

$P(\sigma^2 \mid X) = \text{Inv Gamma}\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$

If you need CR. or Hyp test in real world

Remember $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$

$P(X^* \mid X)$ the next realization, want to know distribution, based on what you've seen in X.

Note: If you knew $X^* \mid \theta, \sigma^2 \sim N(\theta, \sigma^2)$, much easier. comes from model.

$$P(X^* \mid X) = \int_0^\infty \int_{\mathbb{R}} P(X^* \mid \theta, \sigma^2) \underbrace{P(\theta, \sigma^2 \mid X)}_{\text{Norm Inv Gamma}} \, d\theta \, d\sigma^2$$

averaging over what θ, σ^2 look like

$$\propto \iint \left((\sigma^2)^{-\frac{1}{2}} e^{-\frac{1}{2\sigma^2}(X^* - \theta)^2}\right)\left((\sigma^2)^{-\frac{n}{2}-1} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta)^2}\right) d\theta \, d\sigma^2$$

$$\propto \int_0^\infty (\sigma^2)^{-(\frac{n}{2}+1)-1} \underbrace{\int_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}\left((X^* - \theta)^2 + \sum (x_i - \theta)^2\right)}}_{\propto \text{ Normal}} d\theta \, d\sigma^2$$

$\vdots$

$\propto$ Normal. then you can multiply by constant to make into pdf. $= 1$

$$\propto T_{n-1}\left(\bar{x}, s\sqrt{\frac{n+1}{n}}\right)$$

$$\propto N(\theta, \sigma^2) \text{ as } n \text{ gets bigger.}$$

Math 242 : $T_n \to Z \sim N(0,1)$ $[T_n \approx N]$ $(n \geq 30)$

$\lim_{n \to \infty} \frac{n+1}{n} \to 1$

$n \to \infty$ $\bar{X} \to \theta$

$n \to \infty$ $S \to \sigma$

$$= \int_0^\infty \int_{\mathbb{R}} \underbrace{P(X^* \mid \theta, \sigma^2)}_{N(\theta, \sigma^2)} \underbrace{P(\theta \mid X, \sigma^2)}_{N\left(\bar{X}, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)} \underbrace{P(\sigma^2 \mid X)}_{\text{Inv Gamma}\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)} d\theta \, d\sigma^2$$

How do I sample from $P(X^* | X)$?

1. Sample σ_0^2 from InvGamma $\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$
2. Sample θ_0 from $N\left(\bar{X}, \left(\frac{\sigma_0}{\sqrt{n}}\right)^2\right)$
3. Sample X^* from $N(\theta_0, \sigma_0^2)$
4. Repeat steps 1-3 many times and return only $X_1^*, \ldots, X_S^*$.

Do it over but Δ 1 thing.

$X_1, \ldots, X_n | \theta, \sigma^2 \overset{iid}{\sim} N(\theta, \sigma^2)$

$P(\theta, \sigma^2) = P(\theta) P(\sigma^2)$ apriori independence.

$\theta \sim N(\mu_0, \tau^2)$ → used when σ^2 was known

$\sigma^2 \sim \text{InvGamma}\left(\frac{n_0}{2}, \frac{n_0 \sigma_0^2}{2}\right)$ → used when θ was known.

If $\tau^2 = \frac{\sigma^2}{m}$ → everything gets broken. θ, σ^2 dependent ✓.

Consider the case when $\tau^2 \neq \frac{\sigma^2}{m}$, ~~everything gets broken~~.

$$P(\theta, \sigma^2 | X) \propto P(X|\theta, \sigma^2) P(\theta) P(\sigma^2) \propto \left((\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}((n-1)s^2 + n(\bar{x}-\theta)^2)}\right) \left(e^{-\frac{1}{2\tau^2}(\theta-\mu_0)^2}\right) \left((\sigma^2)^{-\left(\frac{n_0}{2}+1\right)} e^{-\frac{n_0\sigma_0^2/2}{\sigma^2}}\right)$$

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

$$\propto \underbrace{(\sigma^2)^{-\frac{n}{2} - \left(\frac{n_0}{2}+1\right)} e^{-\frac{1}{2\sigma^2}((n-1)s^2 + n_0\sigma_0^2 + n\bar{x}^2)}}_{\text{call } A} e^{-\left(\frac{n}{2\sigma^2} + \frac{1}{2\tau^2}\right)\theta^2 + \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\tau^2}\right)\theta} \leftarrow \text{kernel}$$

$$\propto N\left(\underbrace{\frac{\frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\tau^2}}{\frac{n}{2\sigma^2} + \frac{1}{2\tau^2}}}_{\theta_p}, \underbrace{\frac{1}{\frac{n}{2\sigma^2} + \frac{1}{2\tau^2}}}_{\sigma_p^2}\right)$$

so you $\neq$ have to write again.

Putting constants back in:

$$\propto A \sqrt{2\pi\sigma_p^2}\, e^{-\frac{\theta_p^2}{2\sigma_p^2}} N(\theta_p, \sigma_p^2) = (\sigma^2)^{-\frac{n}{2} - \left(\frac{n_0}{2}+1\right)} e^{-\frac{1}{2\sigma^2}((n-1)s^2 + n_0\sigma_0^2 + n\bar{x}^2)} \underbrace{\sqrt{2\sigma \frac{1}{\frac{n}{2\sigma^2} + \frac{1}{2\tau^2}}}\, e^{-\frac{\theta_p^2}{2\sigma_p^2}}}_{} N(\theta_p, \sigma_p^2).$$

Some kernel we've never seen before.

$K(\sigma^2 | x)$

$$= N(\theta_p, \sigma_p^2) K(\sigma^2 | x)$$

$\not\propto$ any known distribution — we're in trouble.

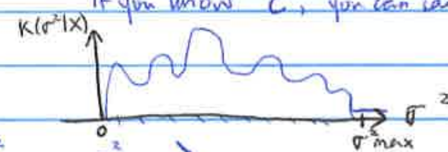
How to sample the posterior $P(\theta, \sigma^2|X)$ (usually garbage)

1. Sample σ_0^2 from $K(\sigma^2|X)$
2. Sample θ_0 from $N\left(\dfrac{\frac{n\bar{x}}{\sigma_0^2}+\frac{\mu_0}{\tau^2}}{\frac{n}{\sigma_0^2}+\frac{1}{\tau^2}}, \dfrac{1}{\frac{n}{\sigma_0^2}+\frac{1}{\tau^2}}\right)$
3. Record $\langle\theta_0, \sigma_0^2\rangle$
4. Repeat steps 1-3 S times.

Remember $P(\sigma^2|X) = cK(\sigma^2|X)$

if you know c, you can sample it

How to sample from $K(\sigma^2|X)$.

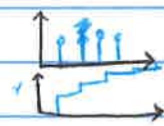


"Step 0"

1. Pick σ^2_{min}, σ^2_{max}, $\Delta\sigma^2$
2. Create grid $G := \langle\sigma^2_{min}, \sigma^2_{min}+\Delta\sigma^2, \sigma^2_{min}+2\Delta\sigma^2, \dots, \sigma^2_{max}\rangle$
3. Compute c. $c \approx \dfrac{1}{\sum_{\sigma^2\in G} K(\sigma^2|X)}$
4. Compute $F(\sigma_0^2|X) \approx \sum_{\{\sigma^2\in G \,:\, \sigma^2 \le \sigma_0^2\}} cK(\sigma^2|X) \quad \forall\ \sigma_0^2 \in G$.

"Step 1"

5. Draw y from $U(0,1)$
6. Compute $\sigma_x^2 \approx F^{-1}(y)$
7. Repeat steps 5 & 6 many times.



to go back to top of pg.

Grid Sampling Disadvantages

1. Numerically Unstable
   a. [Computers have trouble with big numbers & very small numbers]
2. How to pick arbitrary decisions for θ_{min}, θ_{max}, $\Delta\theta$.
3. Imagine $\theta_{min} = 0$, $\theta_{max} = 1$, $\Delta\theta = 10^{-5}$ this means the size ($|G|$) of the grid $= 10^5$. If posterior had 10 dimensions then size of grid $= |G| = 10^{5^{10}} = 10^{50}$ $\Rightarrow$ Impossible!

This really works, in low dimensions if you know the support (to pick θ_{min}, θ_{max}) and you know the shape (to pick $\Delta\theta$).