Lec 16 Math 391 4/27/17

$X_1, \ldots, X_n | \theta, \sigma^2 \overset{iid}{\sim} N(\theta, \sigma^2)$ but θ & σ^2 unknown

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

$P(\theta, \sigma^2 | X) = NInvG$ $\longrightarrow$ Why? You can sample...

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

$P(\sigma^2 | X, \theta) = InvG\left(\frac{n}{2}, \frac{n\hat{\sigma}^2}{2}\right)$

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

$P(\sigma^2 | X) = InvG\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$

very important for inference. Bayesian answer to nuisance params.

CR for θ? Hyp test for θ?

CR for σ^2? ... σ^2?

Next question...

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

$X^* | X \sim$? Use what we got....

Next question... post. pred. distr.

$$P(x^*|X) = \int_{0}^{\infty}\int_{-\infty}^{\infty} P(x^*|\theta,\sigma^2)\, P(\theta,\sigma^2|X)\, d\theta\, d\sigma^2$$

$$\propto \int\int (\sigma^2)^{-\frac{1}{2}} e^{-\frac{1}{2\sigma^2}(x^*-\theta)^2} (\sigma^2)^{-\frac{n}{2}-1} e^{-\frac{1}{2\sigma^2}\Sigma(x_i-\theta)^2}\, d\theta$$

$$= \int (\sigma^2)^{-\left(\frac{n+1}{2}\right)-1} \int e^{-\frac{1}{2\sigma^2}\left((x^*-\theta)^2 + \Sigma(x_i-\theta)^2\right)}\, d\theta\, d\sigma^2$$

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

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

$$= \underbrace{\left(x^{*2} + n\bar{x}^2 + (n-1)s^2\right)}_{a} - 2(x^* + n\bar{x})\theta + (n+1)\theta^2$$

$$= \int (\sigma^2)^{-\left(\frac{n+1}{2}\right)-1} e^{-\frac{a}{2\sigma^2}} \int \underbrace{e^{\frac{x^*+n\bar{x}}{\sigma^2}\theta}\, e^{-\frac{n+1}{2\sigma^2}\theta^2}}_{\text{kernel of normal}}\, d\theta\, d\sigma^2$$

$$-\frac{1}{2v^2} = -\frac{n+1}{2\sigma^2} \Rightarrow v^2 = \frac{\sigma^2}{n+1}$$

$$\frac{c}{v^2} = \frac{x^*+n\bar{x}}{\sigma^2} \Rightarrow c = \frac{(x^*+n\bar{x})}{\sigma^2}\cdot\frac{\sigma^2}{n+1} = \frac{x^*+n\bar{x}}{n+1}$$

$$f(c,v^2) = \frac{1}{\sqrt{2\pi v^2}} e^{-\frac{1}{2v^2}(\theta-c)^2} = \frac{1}{\sqrt{2\pi v^2}} e^{-\frac{1}{2v^2}\theta^2} e^{\frac{c}{v^2}\theta} e^{\frac{c^2}{2v^2}}$$

$$\int e^{\frac{c}{v^2}\theta} e^{-\frac{1}{2v^2}\theta^2} e^{\frac{c^2}{2v^2}} e^{-\frac{c^2}{2v^2}} \frac{1}{\sqrt{2\pi v^2}} \sqrt{2\pi v^2}$$

$$\propto T_{n-1}\left(\bar{x}, \sqrt{s^2 \frac{n+1}{n}}\right)$$

If n large $T_{n-1} \approx N$, $\frac{n+1}{n} \approx 1 \Rightarrow x^*|x \approx N(\bar{x}, s^2)$

which makes sense...

How to draw from this distr?

$$p(x^*|x) = \iint p(x^*|\theta,\sigma^2)\, p(\theta|x,\sigma^2)\, p(\sigma^2|x)\, d\sigma^2\, d\theta$$

$p(x^*|\theta,\sigma^2) \to N(\theta,\sigma^2)$, $p(\theta|x,\sigma^2) \to N\left(\bar{x}, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$, $p(\sigma^2|x) \to \text{InvGamma}\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$

Step 1: Draw σ^2 from $\text{InvGamma}\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$

Step 2: Draw θ from $N\left(\bar{x}, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$

Step 3: Draw x^* from $N(\theta,\sigma^2)$

Return only x^*

Remember...

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

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

$$\Rightarrow p(\theta,\sigma^2|x) = \text{NormInvGamma}$$

Also... inverse gamma normal-posterior...

Also, if $p(\theta|\sigma^2) = N\left(\mu_0, \left(\frac{\sigma}{\sqrt{n_0}}\right)^2\right)$, $p(\sigma^2) = \text{InvGamma}\left(\frac{n_0}{2}, \frac{n_0\sigma_0^2}{2}\right)$

What if

$P(\theta) = N(\mu_0, \tau^2)$

$P(\sigma^2) = \text{InvGamma}\left(\frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2}\right)$

s.t. $\tau^2 \neq \frac{\sigma^2}{\nu_0}$

Note $P(\theta, \sigma^2) = P(\theta) P(\sigma^2)$ i.e. a priori independent

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

$$\propto (\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\left((n-1)s^2 + n(\bar{X}-\theta)^2\right)} e^{-\frac{1}{2\tau^2}(\theta-\mu_0)^2} (\sigma^2)^{-\left(\frac{\nu_0}{2}+1\right)} e^{-\frac{\nu_0\sigma_0^2/2}{\sigma^2}}$$

$$= (\sigma^2)^{-\frac{n}{2}-\left(\frac{\nu_0}{2}+1\right)} e^{-\frac{1}{2\sigma^2}\left((n-1)s^2 + \nu_0\sigma_0^2\right)} e^{-\frac{n}{2\sigma^2}(\bar{X}-\theta)^2 - \frac{1}{2\tau^2}(\theta-\mu_0)^2}$$

$$-\frac{n\bar{X}^2}{2\sigma^2} + \frac{n\bar{X}\theta}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} - \frac{\theta^2}{2\tau^2} + \frac{\theta\mu_0}{\tau^2} - \frac{\mu_0^2}{2\tau^2}$$

$$\propto (\sigma^2)^{-\frac{n}{2}-\left(\frac{\nu_0}{2}+1\right)} e^{-\frac{1}{2\sigma^2}\left((n-1)s^2 + \nu_0\sigma_0^2 + n\bar{X}^2\right)} \cdot \underbrace{e^{-\left(\frac{n}{2\sigma^2}+\frac{1}{2\tau^2}\right)\theta^2 + \left(\frac{n\bar{X}}{\sigma^2}+\frac{\mu_0}{\tau^2}\right)\theta}}_{\propto N(\theta_p, \sigma_p^2)}$$

$$= (\sigma^2)^{-\frac{n}{2}-\left(\frac{\nu_0}{2}+1\right)} e^{-\frac{1}{2\sigma^2}\left((n-1)s^2+\nu_0\sigma_0^2+n\bar{x}^2\right)} \cdot \sqrt{2\pi\sigma_p^2}\; e^{-\frac{\theta_p^2}{2\sigma_p^2}} \qquad N(\theta_p, \sigma_p^2)$$

$$\underbrace{}_{} \qquad \propto \qquad \frac{1}{\sqrt{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}} \quad e^{-\frac{1}{2}\frac{\left(\frac{n\bar{x}}{\sigma^2}+\frac{\mu_0}{\tau^2}\right)^2}{\left(\frac{n}{\sigma^2}+\frac{1}{\tau^2}\right)^3}} \qquad \frac{1}{\sqrt{2\pi\sigma_p^2}} e^{-\frac{1}{2\sigma_p^2}(\theta-\theta_p)^2}$$

$K(\sigma^2|x)$ $\not\propto$ InvGamm

nor anything else known.

But it is a kernel of some r.v.

$\Rightarrow$ "Semi-conjugate" Model

$\Rightarrow$ only conjugate for posteriors of one parameter i.e. $P(\theta|X,\sigma^2)$, $P(\sigma^2|X,\theta)$ but not $P(\theta,\sigma^2|x)$!

What if we want to sample?

Step 1 : sample σ_0^2 from $k(\sigma^2|x)$

Step 2 : sample θ_0 from $N\left(\theta_p, \sigma_p^2 = \frac{1}{\frac{n}{\sigma_0^2}+\frac{1}{\tau^2}}\right)$

Step 3 : record $\langle \theta_0, \sigma_0^2 \rangle$

How to do step 1???

Recall $P(\sigma^2|x) = c\, K(\sigma^2|x)$

Create grid, set $\sigma^2_{min}, \sigma^2_{max}, \Delta\sigma^2$

$G = \{\sigma^2_{min}, \sigma^2_{min}+\Delta x, \sigma^2_{min}+2\Delta x, \ldots, \sigma^2_{max}\}$

$c \approx \dfrac{1}{\sum_{\sigma^2\in G} K(\sigma^2|x)}$ $\Rightarrow P(\sigma^2|x) \approx c\,k(\sigma^2|x)$

$\Rightarrow F(\sigma_0^2|x) \approx \sum_{\{\sigma^2\in G:\, \sigma^2<\sigma_0^2\}} c\,k(\sigma^2|x)$

Grid Sampling

Now draw y from $U(0,1)$. Compute $\sigma_0^2 = \min_{\sigma^2\in G} F(\sigma^2) \geq y$

Grid Sampling Disadvantages

① Numerically unstable.

Computers have minimum values of #'s / max. values of #'s.

② How to pick θ_{min}, θ_{max}, $\Delta\theta$?

Bad decision for $\theta_{min}, \theta_{max} \Rightarrow$ You miss a part of the support of the parameter!

Bad decision for $\Delta\theta \Rightarrow$ Bad resolution $\Rightarrow$ non-reliable samples

③ Let's say $\theta_{min} = 0$, $\theta_{max} = 1$, $\Delta\theta = 0.0001$, $|G| = 10{,}000 = 10^5$

What if θ had 10 dimensions? $\Rightarrow |G| = 10^{5^{10}} = 10^{50} \Rightarrow$ IMPOSSIBLE for a computer!

$\Rightarrow$ Grid sampling only good in low dimensions if you know the effective support of θ (i.e. where most of the support lies) and if you know the shape so you can pick a reasonable $\Delta\theta$.

It would be nice to fix these problems with a new method. We will do this LAST. Now onto something different...