Lecture 14 3/30/16

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

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

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

$\frac{1}{n-1}\sum(x_i - \bar{x})^2$ unbiased estimator for variance "Sample var"

N-InvGamma distr

$$P(\sigma^2|X) = \int_{Supp(\theta)} P(\theta, \sigma^2|X)\, d\theta = \int_{\mathbb{R}} (\sigma^2)^{-(\frac{n}{2}+1)} e^{-\frac{1}{2\sigma^2}\left((n-1)s^2 + n(\bar{x}-\theta)^2\right)}\, d\theta$$

why

$$= (\sigma^2)^{-(\frac{n}{2}+1)} e^{-\frac{1}{2\sigma^2}(n-1)s^2} \underbrace{\int_{\mathbb{R}} \underbrace{e^{-\frac{n}{2\sigma^2}(\bar{x}-\theta)^2}}_{\text{Kernel of Norm}}\, d\theta}_{\propto 1}$$

$$f(x|z) = \int_Y f(x,y|z)\, dy = \int_Y f(x|y,z)\, f(y|z)\, dy$$

"marginalization"

if distr is $P(y)$ then integral $\exists$ AVG!

$\propto$ InvGamma$\left(\frac{n}{2}, \frac{(n-1)s^2}{2}\right)$

Not required

Prereq's $Z_0, Z_1, \ldots, Z_\nu \stackrel{iid}{\sim} N(0,1)$

$$\sum_{i=1}^{\nu} Z_i^2 \sim \chi^2_\nu := \frac{2^{-\nu/2}}{\Gamma(\nu/2)} \theta^{\frac{\nu}{2}-1} e^{-\frac{\theta}{2}} \stackrel{d}{=} \text{Gamma}\left(\frac{\nu}{2}, \frac{1}{2}\right)$$

free variable

$$\frac{Z_0}{\sqrt{\sum_{i=1}^{\nu} Z_i^2}} \sim T_\nu := \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\nu\pi}} \left(1 + \frac{1}{\nu}\theta^2\right)^{-\frac{\nu+1}{2}}$$

$\bar{X} \sim T_{n-1}\left(\mu, \sqrt{\frac{s^2}{n}}\right) :=$ Student PDF

centered and scaled T

Classic results (Math 242)

5 lectures to prove these

$\frac{n-1}{\sigma^2} s^2 \sim \chi^2_{n-1}$, $\frac{\bar{X} - \mu}{\sqrt{\frac{s^2}{n}}} \sim T_{n-1}$ "Standard"

two distr's and two results

$T_1 \stackrel{d}{=}$ Std. Cauchy

$s^2 \to \sigma^2$ so $T_{n-1} \to Z$ Normal with fatter tails

$$\Rightarrow P(\theta, \sigma^2 | X) = \underbrace{N\left(\bar{X}, \frac{\sigma^2}{n}\right)}_{P(\theta|\sigma^2, X)} \underbrace{\text{InvGamma}\left(\frac{n}{2}, \frac{(n-1)S^2}{2}\right)}_{P(\sigma^2|X)}$$

Sometimes called N-InvGamma distr.

Expectation? Variance? Not covered...

How to sample?

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

Step 2: Draw θ from $N\left(\bar{X}, \frac{\sigma^2}{n}\right)$ using σ^2 from Step 1

Repeat many, many times

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

How to make it informative?? (General model can incorporate prior info)

we assumed $P(\sigma^2 | \theta) = P(\sigma^2)$ before

Same as had $U(0,1)$ but want $\text{Beta}(\alpha, \beta)$

or $\text{Gamma}(1,0)$ but want $\text{Gamma}(\alpha, \beta)$

or $N(0, \infty)$ but want $N(\mu_0, \tau^2)$

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

$\Uparrow$

if $\theta|\sigma^2 \sim N\left(\mu_0, \frac{\sigma^2}{m}\right)$

$\tau^2 = f(\sigma^2)$

why??

What does prior mean? m measurements

Look at the above

It seems prior of θ should be dependent on prior of σ^2

τ^2 is a function of σ^2 now directly

$$\Rightarrow \theta | \sigma^2 \sim N\left(\mu_0, \frac{\sigma^2}{m}\right), \quad \sigma^2 \sim \text{InvGamma}\left(\frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2}\right)$$

$\Rightarrow \theta, \sigma^2 \sim \text{N InvGamma}$ ... conjugate for Normal with params θ, σ^2

For HW...

$$p(\theta, \sigma^2 | X) = N\left(\frac{n}{n+m}\bar{X} + \frac{m}{n+m}\mu_0, \ \frac{\sigma^2}{n+m}\right)$$

$$\cdot \text{InvGamma}\left(\frac{\nu_0 + n}{2}, \ \frac{\nu_0 \sigma_0^2 + (n-1)S^2 + \frac{mn}{m+n}(\bar{X}-\mu_0)^2}{2}\right)$$

To draw samples

Step 1: σ^2 draw InvG

Step 2: θ from normal

using σ^2 from step 1

HW for everyone — prove my conjecture...

$\Rightarrow$ multiparameter Bayesian models are really hard!!

θ, σ^2 unknown but only care about θ...

WANT:

$p(\theta | X)$ How does this differ from $p(\theta | X, \sigma^2)$ which we did before??

we don't know this

...nd θ, but σ^2 "gets in the way"! σ^2 is called a "nuisance parameter". What to do?

Recall $f(x) = \int_{\text{supp}[Y]} f(x, y)\,dy$, $f(x|z) = \int_{\text{supp}[Y]} f(x, y|z)\,dy$

$\Rightarrow p(\theta | X) = \int_{\text{Supp}[\sigma^2]} p(\theta, \sigma^2 | X)\,d\sigma^2$ margin out σ^2... Avg. your idea about θ over all possible σ^2's.

Use non-informative prior function first... 4

$$p(\theta|x) \propto \int_0^{\infty} (\sigma^2)^{-\frac{n}{2}-1} e^{-\frac{1}{2\sigma^2}\left((n-1)s^2 + n(\bar{x}-\theta)^2\right)} d\sigma^2$$

let $A := (n-1)s^2 + n(\bar{x}-\theta)^2$ ← not a function of σ^2

$$t := \frac{A}{2\sigma^2} \Rightarrow \sigma^2 = \frac{A}{2t} \Rightarrow \frac{d\sigma^2}{dt} = -\frac{A}{2t^2}$$

$$\Rightarrow d\sigma^2 = -\frac{A}{2t^2}dt$$

$$\sigma^2 \in (0,\infty) \Rightarrow t \in (0,\infty)$$

$$= \int_0^{\infty} \left(\frac{A}{2t}\right)^{-\frac{n}{2}-1} e^{-t}\left(-\frac{A}{2t^2}dt\right)$$

$$= -\int_0^{\infty} \frac{A^{-\frac{n}{2}}}{2^{-\frac{n}{2}-1}\, t^{-\frac{n}{2}-1+2}} e^{-t}dt$$

$$t^{-\left(-\frac{n}{2}+1\right)} = t^{\frac{n}{2}-1} = t^{\frac{n-2}{2}}$$

$$\propto A^{-\frac{n}{2}} \underbrace{\int_0^{\infty} t^{\frac{n-2}{2}} e^{-t}dt}_{\text{Kernel of gamma} \Rightarrow \text{constant}}$$

$$\propto A^{-\frac{n}{2}}$$

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

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

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

the T is the overdispersed Normal when σ^2 is allowed to vary...

non-central scale Student t distr.

$$X \sim T(\nu, \mu, \sigma^2)$$

$$:= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{2\pi\sigma^2}} \left(1 + \frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}}$$

$$\propto \left(1 + \frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}}$$

Traditionally, the non-central scaled t distr is centered and scaled (standardized) to yield:

$$\Rightarrow X \sim T(\nu, \mu, \sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \sim T_\nu$$

Student's T distr.

Note: $T_\nu \overset{d}{=} \frac{Z}{\sqrt{X_\nu^2}}$ where $Z \sim N(0,1)$

X^2 is $\chi^2_\nu := \sum_{i=1}^{\nu} Q_i^2$, $Q_i \overset{iid}{\sim} N(0,1)$

this is the test statistic in classic statistics for hypothesis tests of μ when σ^2 unknown

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

By a similar exercise... If $P(\theta|\sigma^2) = N\left(\mu_0, \frac{\sigma^2}{m}\right)$ and $P(\sigma^2) = \text{InvGamma}\left(\frac{\nu_0}{2}, \frac{\nu_0\sigma_0^2}{2}\right)$

$$\Rightarrow P\left(\frac{\theta - \mu_1}{\frac{\sigma_1}{\sqrt{m}}} \,\middle|\, X\right) = T_{\nu_0 + n}$$

where $\left.\begin{array}{l} \mu_1 = \frac{n}{n+m}\bar{x} + \frac{m}{n+m}\mu_0 \\ \sigma_1 = \sqrt{\frac{\nu_0\sigma_0^2}{n+m} + \frac{mn}{(n+m)^2}(n-1)s^2} \end{array}\right\}$ I think..... (HW)

$$= \int_{\mathbb{R}} P(\theta, \sigma^2 | X)\, d\theta$$

$$P(\sigma^2 | X) = \text{InvGamma}\left(\frac{\nu_0 + n}{2}, \frac{1}{2}\left(\nu_0 \sigma_0^2 + (n-1)s^2 + \frac{nm}{n+m}(\bar{x} - \mu_0)^2\right)\right)$$

⇒ McGrayne ... Bayesian Models very computational!

This is only $N(\theta, \sigma^2)$ with conj. prior!!!

What about bigger models?? WORSE!!!

Need better way....

Under conj. priors...

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

$$= \int\int \underbrace{P(X^* | \theta, \sigma^2)}_{N(\theta, \sigma^2)} \cdot \underbrace{P(\theta | X, \sigma^2)}_{N\left(\frac{n}{n+m}\bar{x} + \frac{m}{n+m}\mu_0, \frac{\sigma^2}{n+m}\right)} \cdot \underbrace{P(\sigma^2 | X)}_{\text{InvGamma}\left(\frac{\nu_0 + n}{2}, \frac{\nu_0 \sigma_0^2 + ns^2}{2}\right)}\, d\sigma^2\, d\theta$$

$$\vdots$$

$$= T\left(n-1, \bar{x}, s\sqrt{1 + \tfrac{1}{n}}\right)$$ Super Herculean effort! (EC)

OR: How to draw samples?

Step 1: Draw σ^2 from InvGam.

Step 2: Draw θ from Normal using σ^2 from Step 1.

Step 3: Draw X^* from Normal using θ, σ^2 from Step 1, 2.

DO EX NOW