Lec 18 Math 341 4/9/19 Start on third page of previous notes

1

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

Assume σ^2 known

Conj. prior $P(\theta) = N(\mu_0, \tau^2)$ — Laplace: $N(0, \infty)$; Haldane: ''; Jeffreys: ''

Posterior: $$P(\theta \mid x) = N\left(\underbrace{\frac{\frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}}_{\theta_p}, \underbrace{\frac{1}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}}_{\sigma_p^2}\right)$$

$\hat{\theta}_{MMSE} = \hat{\theta}_{MAP} = \hat{\theta}_{MMAE} = \theta_p$, $\hat{\theta}_{MLE} = \bar{X}$, shrinkage: $\rho = \frac{\sigma^2}{n\tau^2 + \sigma^2}$

CR's: use qnorm!, Hyp tests: use pnorm!

Posterior predictive: $P(X^* \mid X, \sigma^2) = N(\theta_p, \sigma^2 + \sigma_p^2)$

If conj. prior was parameterized using σ^2, $\tau^2 = \frac{\sigma^2}{n_0}$

$$P(\theta \mid \sigma^2) = N\left(\mu_0, \frac{\sigma^2}{n_0}\right)$$

Posterior: $$P(\theta \mid x) = N\left(\underbrace{\frac{n\bar{x} + n_0\mu_0}{n + n_0}}_{\theta_p}, \underbrace{\frac{\sigma^2}{n + n_0}}_{\sigma_p^2}\right)$$

shrinkage: $\rho = \frac{n_0}{n + n_0}$

Posterior predictive: $$P(X^* \mid X, \sigma^2) = N\left(\theta_p, \sigma^2 \frac{n + n_0 + 1}{n + n_0}\right)$$

Assume θ known

Conj prior: $P(\sigma^2) = \text{InvGamma}\left(\frac{n_0}{2}, \frac{n_0\sigma_0^2}{2}\right)$ — Laplace: InvGamma$(-1, 0)$; Haldane: InvGamma$(0, 0)$; Jeffreys: InvGamma$(0, 0)$

Posterior: $$P(\sigma^2 \mid X, \theta) = \text{InvGamma}\left(\frac{n + n_0}{2}, \frac{n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}{2}\right)$$

shrinkage: $\rho = \frac{n_0 - 2}{n + n_0 - 2}$

$\hat{\sigma}^2_{MMSE} = \frac{n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}{n + n_0 - 2}$, $\hat{\sigma}^2_{MAP} = \frac{n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}{n + n_0 + 2}$, $\hat{\sigma}^2_{MMAE} = \text{qinvgamma}\left(0.5, \frac{n + n_0}{2}, \frac{n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}{2}\right)$

post. predictive distr.: $$P(X^* \mid X, \theta) = T_{n + n_0}\left(\theta, \frac{n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}{n + n_0}\right)$$

Midterm 2 ~~!!!~~

FINAL!!!

the real problem!

F: $X_1, \dots, X_n \mid \theta, \sigma^2 \stackrel{iid}{\sim} N(\theta, \sigma^2)$ where both θ, σ^2 unknown.

We want inference for both at the same time. Never seen this!

$$\underbrace{P(\theta, \sigma^2 \mid X)}_{\text{2-dim posterior}} \propto P(X \mid \theta, \sigma^2) \underbrace{P(\theta, \sigma^2)}_{\text{2-d prior}}$$

$$= (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum(X_i - \theta)^2}$$

$$\propto (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2} \stackrel{?}{\propto} \text{InvGamma?}$$ No! θ is a free variable.

this is actually called the NormalInvGamma Distr.

But to get it into canonical form, we need more manipulation.

Note: $s^2 := \frac{1}{n-1}\sum(X_i - \bar{X})^2$

trick: $\sum(X_i - \theta)^2 = \sum((X_i - \bar{X}) + (\bar{X} - \theta))^2 = \sum(X_i - \bar{X})^2 + 2\sum(X_i - \bar{X})(\bar{X} - \theta) + \sum(\bar{X} - \theta)^2$

$$= (n-1)s^2 + 2\left(\sum X_i\bar{X} - \bar{X}^2 - X_i\theta + \bar{X}\theta\right) + n(\bar{X} - \theta)^2$$

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

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

$$\Rightarrow = \underbrace{(\sigma^2)^{-(n/2+1)-1} e^{-\frac{(n-1)s^2/2}{\sigma^2}}}_{\text{"inv gamma" kernel}} \underbrace{e^{-\frac{n}{2\sigma^2}(\theta - \bar{X})^2}}_{\text{"normal kernel"}} \propto \text{NormInvGamma}\left(\mu = \bar{X}, \lambda = n, \alpha = \frac{n}{2}+1, \beta = \frac{(n-1)s^2}{2}\right)$$

$\propto \underbrace{P(\sigma^2 \mid X)}_{\text{Inv.}} \underbrace{P(\theta \mid \sigma^2, X)}_{\text{Norm}}$

4 parameters!!

$\Rightarrow$ $P(\theta, \sigma^2) = \text{NormInvGamma}$ is the conjugate prior.

for the purposes of this class, we will only consider the uninformative Jeffrey's prior which can be shown to be:

$$P(\theta, \sigma^2) = \text{NormInvGamma}(0, \infty, 0, 0) \propto \frac{1}{\sigma^2} = P_J(\theta, \sigma^2)$$

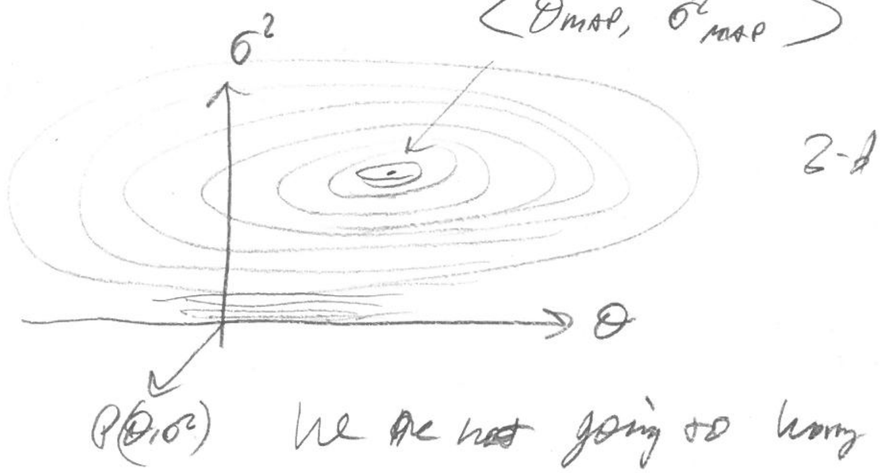
this is the Jeffreys prior for the normal with σ^2 known times the " " " " " " " " " " " " θ known.

$$P_J(\theta, \sigma^2) = P_J(\theta) P_J(\sigma^2) \propto (1)\left(\frac{1}{\sigma^2}\right) = \frac{1}{\sigma^2}$$

$$\Rightarrow P(\theta, \sigma^2 | x) \propto P(x | \theta, \sigma^2) P_J(\theta, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}-1} e^{-\frac{(n-1)s^2/2}{\sigma^2}} e^{-\frac{n}{2\sigma^2}(\theta - \bar{x})^2}$$

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

What does this look like?



2-d density as a contour

We are not going to worry about point est's, CR, Hyp. Tests right now for this distr., but we will do marginal distr's and posterior predictive distr's

The usual case is ~~we~~ just care about θ. σ^2 is then a nuisance parameter. Thus we are interested in only $P(\theta|X)$. We can "get rid of σ^2" by marginalizing it out!

$$P(\theta|x) = \int_0^\infty P(\theta,\sigma^2|x)\, d\sigma^2$$

Let's solve this assuming $P(\theta,\sigma^2) \propto \frac{1}{\sigma^2}$

$$\Rightarrow P(\theta,\sigma^2|x) = \text{NormInvGamma}\left(\bar{x}, n, \frac{n}{2}, \frac{(n-1)s^2}{2}\right) \propto (\sigma^2)^{-\frac{n}{2}-1} e^{-\frac{(n-1)s^2/2}{\sigma^2}} e^{-\frac{n}{2\sigma^2}(\theta-\bar{x})^2}$$

$$\Rightarrow P(\theta|x) \propto \int_0^\infty \underbrace{(\sigma^2)^{-\frac{n}{2}-1} e^{-\frac{(n-1)s^2/2 + n(\theta-\bar{x})^2/2}{\sigma^2}}}_{\propto \text{InvGamma}\left(\frac{n}{2}, \frac{(n-1)s^2 + n(\theta-\bar{x})^2}{2}\right)} d\sigma^2$$

$\frac{n}{2} = \alpha$, $\frac{(n-1)s^2 + n(\theta-\bar{x})^2}{2} = \beta$

$$c = \frac{\beta^\alpha}{\Gamma(\alpha)}$$

$$\Rightarrow P(\theta|x) \propto \frac{\Gamma(\alpha)}{\beta^\alpha} \underbrace{\int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} e^{-\frac{\beta}{\sigma^2}} d\sigma^2}_{=1}$$

we kinda seen this before....

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

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

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

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

i.e. the shifted and scaled Student's T or "non-std T"!