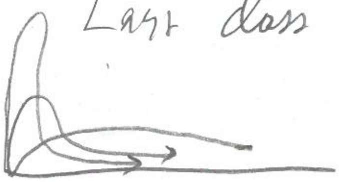
Lec 17 Math 341 4/4/19

F: $X_1,\dots,X_n \mid \theta,\sigma^2 \overset{iid}{\sim} N(\theta,\sigma^2)$ with θ known and we want inference for σ^2.

Last class we derived the conjugate prior as:



Interpretation of hyperparameters

pseudoobs's (strength)

variance of pseudoobs's.

$$P(\sigma^2) = \text{InvGamma}\left(\frac{n_0}{2}, \frac{n_0\sigma_0^2}{2}\right) \propto (\sigma^2)^{-\frac{n_0}{2}-1} e^{-\frac{n_0\sigma_0^2/2}{\sigma^2}}$$

prior exp: $E(\sigma^2) = \dfrac{\frac{n_0\sigma_0^2}{2}}{\frac{n_0}{2}-1} = \dfrac{n_0}{n_0-2}\sigma_0^2$

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

Pseudocount Interpretation:

n_0: # pseudoobservations (strength)

σ_0^2: variance of the pseudoobservations

Pt. Est.

$\Rightarrow \hat\sigma^2_{MMSE} = \dfrac{n\hat\sigma^2_{MLE}+n_0\sigma_0^2}{n+n_0-2}$

$\hat\sigma^2_{MAP} = \dfrac{n\hat\sigma^2_{MLE}+n_0\sigma_0^2}{n+n_0+2}$

$\hat\sigma^2_{MMAE} = \text{QInvGamma}\left(0.5, \dfrac{n+n_0}{2}, \dfrac{n\hat\sigma^2_{MLE}+n_0\sigma_0^2}{2}\right)$

CR's & Hypothesis Tests... same!!

Shrinkage

$$\hat\sigma^2_{MMSE} = \frac{n\hat\sigma^2_{MLE}}{n+n_0-2} + \frac{n_0\sigma_0^2}{n+n_0-2}\cdot\frac{n_0-2}{n_0-2} = \underbrace{\frac{n}{n+n_0-2}}_{1-\rho}\hat\sigma^2_{MLE} + \underbrace{\frac{n_0-2}{n+n_0-2}}_{\rho}E(\sigma^2)$$

Uninformative Priors

① Laplace $P(\sigma^2 \mid \theta) \propto 1$

$$P(\sigma^2 \mid \theta,X) \propto P(X \mid \sigma^2,\theta) \propto (\sigma^2)^{-(n/2-1)-1} e^{-\frac{n\hat\sigma^2_{MLE}/2}{\sigma^2}} \propto \text{InvGamma}\left(\frac{n-2}{2}, \frac{n\hat\sigma^2_{MLE}}{2}\right)$$

only proper if $n \ge 3$

$$\Rightarrow P(\sigma^2 \mid \theta) = \text{InvGamma}(-1, 0) \propto \frac{1}{(\sigma^2)^0}$$

i.e. improper

$n_0 = -2$ $\quad \sigma_0^2 = 0$

② Haldane Prior

$$\propto \frac{1}{\sigma^2}$$

$$n_0 = 0 \Rightarrow p(\sigma^2|\theta) = \text{InvGamma}(0,0) \Rightarrow p(\sigma^2|\theta,x) = \text{InvGamma}\left(\frac{n}{2}, \frac{n\hat{\sigma}^2_{MLE}}{2}\right)$$

(improper)

proper only if $n \geq 1$. ~~If~~ ??

~~If $n=1 \Rightarrow \hat{\sigma}^2_{MLE} = 0$~~

③ Jeffreys Prior

$$l'(\sigma^2; X, \theta) = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum(X_i-\theta)^2$$

$$-l''(\sigma^2; X, \theta) = -\left(\frac{n}{2}\frac{1}{(\sigma^2)^2} + -2\left(\frac{1}{2}\right)\frac{1}{(\sigma^2)^3}\sum(X_i-\theta)^2\right)$$

$$= -\frac{n}{2(\sigma^2)^2} + \frac{\sum(X_i-\theta)^2}{(\sigma^2)^3}$$

$$I(\sigma^2) = E[-l''] = E\left[\frac{-n}{2(\sigma^2)^2} + \frac{\sum(X_i-\theta)^2}{(\sigma^2)^3}\right] = -\frac{n}{2(\sigma^2)^2} + \frac{n\,E[(X_1-\theta)^2]}{2(\sigma^2)^3} = n\left(-\frac{1}{2(\sigma^2)^2} + \frac{1}{2(\sigma^2)^2}\right)$$

$$= \frac{n}{2}(\sigma^2)^{-2}$$

$$E(X_1-\theta)^2 = E(X_1^2 + 2X_1\theta + \theta^2) = E(X_1^2) - 2\theta^2 + \theta^2 = \sigma^2 + \theta^2 - 2\theta^2 + \theta^2$$

$$Var(X) = E(X)^2 - E(X)^2 \Rightarrow E(X^2) = \sigma^2 + \theta^2$$

$$p_J(\sigma^2) \propto \sqrt{I(\sigma^2)} = \sqrt{\frac{n(\sigma^2)^{-2}}{2}} \propto (\sigma^2)^{-1} = (\sigma^2)^{-0-1} \propto \text{InvGamma}(0,0) \propto \frac{1}{\sigma^2}$$

Posterior Predictive Distr.

$$p(X^*|X,\theta) = \int_0^\infty p(X^*|\theta,\sigma^2)\, p(\sigma^2|X,\theta)\, d\sigma^2 = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(X^*-\theta)^2} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)}(\sigma^2)^{-\alpha-1} e^{-\frac{\beta}{\sigma^2}}\, d\sigma^2$$

where $\beta = \frac{n\hat{\sigma}^2 + n_0\sigma_0^2}{2}$, $\alpha = \frac{n+n_0}{2}$

$$\propto \int_0^\infty (\sigma^2)^{-1/2} e^{-\frac{(x^*-\theta)^2/2}{\sigma^2}} (\sigma^2)^{-\alpha-1} e^{-\frac{\beta}{\sigma^2}} \, d\sigma^2$$

$$= \int_0^\infty (\sigma^2)^{-\overbrace{(\alpha+\frac{1}{2})}^{\alpha'}-1} e^{-\frac{\overbrace{(x^*-\theta)^2/2+\beta}^{\beta'}}{\sigma^2}} \, d\sigma^2$$

$$= \frac{\Gamma(\alpha')}{\beta'^{\alpha'}} \overset{\text{normalization constant inverse}}{=} \Gamma\left(\frac{n+n_0+1}{2}\right) \left(\frac{\overbrace{n\hat{\sigma}^2+n_0\sigma_0^2}^{a} + (x^*-\theta)^2}{2}\right)^{-\frac{\overbrace{n+n_0}^{\nu}+1}{2}}$$

$$\propto \left(\frac{a+(x^*-\theta)^2}{2}\right)^{-\frac{\nu+1}{2}} = \left(\frac{2}{a}\right)^{-\frac{\nu+1}{2}} \left(\frac{2}{a}\left(\frac{a+(x^*-\theta)^2}{2}\right)\right)^{-\frac{\nu+1}{2}}$$

$$\propto \left(1+\frac{(x^*-\theta)^2}{a}\right)^{-\frac{\nu+1}{2}} = \left(1+\frac{1}{\nu}\frac{(x^*-\theta)^2}{\frac{a}{\nu}}\right)^{-\frac{\nu+1}{2}} \propto T_\nu\left(\theta, \frac{a}{\nu}\right)$$

shifted & scaled student T

we will define this formally soon

$$= T_{n+n_0}\left(\overset{\text{known}}{\theta}, \frac{n\hat{\sigma}^2+n_0\sigma_0^2}{n+n_0}\right)$$

Mistakes

Review: cheat sheet

$\mathcal{F}: X|\theta,n \sim \text{Bin}(n,\theta)$, $P(\theta) = \text{Beta}(\alpha,\beta)$ $\Rightarrow P(x^*|X,n) = \text{BetaBin}(n^*, X+\alpha, n-X+\beta)$

$\mathcal{F}: X_1,\ldots,X_n|\theta \overset{iid}{\sim} \text{Poisson}(\theta)$

Conj prior: $P(\theta) = \text{Gamma}(\alpha,\beta)$

Laplace: $\text{Gamma}(1,0)$
Haldane: $\text{Gamma}(0,0)$
Jeffreys: $\text{Gamma}(\frac{1}{2},0)$

Posterior: $P(\theta|x) = \text{Gamma}(\Sigma x_i+\alpha, n+\beta)$

$\hat{\theta}_{MMSE} = \frac{\Sigma x_i+\alpha}{n+\beta}$, $\hat{\theta}_{MAP} = \frac{\Sigma x_i+\alpha-1}{n+\beta}$, $\hat{\theta}_{MMAE} = \text{qgamma}(0.5, \Sigma x_i+\alpha, n+\beta)$, $\hat{\theta}_{MLE} = \bar{x}$

CR's: use qgamma approp! Hyp tests: use pgamma appropriately! Shrinkage: $\rho = \frac{\beta}{n+\beta}$

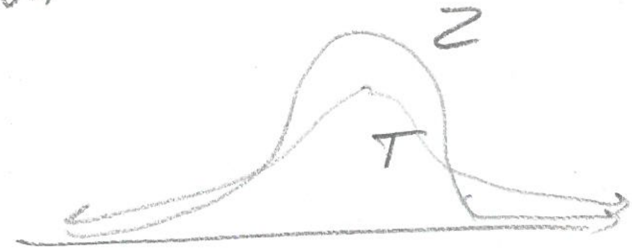
Post. pred: $P(x^*|x) = \text{NegBin}\left(\Sigma x_i+\alpha, \frac{n+\beta}{n+\beta+1}\right)$

What is the T distr? Back to 281... Let $X_1 \ldots X_n | \theta, \sigma^2 \stackrel{iid}{\sim} N(\theta, \sigma^2)$

$$\Rightarrow \frac{\bar{X} - \theta}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$ Basis for "z-test"

But what if σ unknown? We know $s \approx \sigma$ $\quad s = \sqrt{\frac{1}{n-1}\sum(x_i - \bar{x})^2}$

$$\frac{\bar{X} - \theta}{\frac{s}{\sqrt{n}}} \sim T_{n-1}(0,1) = T_{n-1}$$ ← Shorthand: the standard Student's T distr.



Looks the same but has thicker tails.

Recall if $P(\theta) \propto 1$

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

$$P(\theta | x) = T_{n-1}\left(\bar{x}, \frac{s}{\sqrt{n}}\right)$$ Same thing with thicker tails!

FYI

If $V \sim T_n := \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)}\left(1 + \frac{v^2}{n}\right)^{-\frac{n+1}{2}} = T_n(0,1) = T_n$ (standard T)

(non-scaled T)

$W = \sigma V + \mu = t(v) \Rightarrow v = \frac{w - \mu}{\sigma} = t^{-1}(w)$ $\quad = T_n(\mu, \sigma)$

$$f_W(w) = f_V(t^{-1}(w))\left|\frac{d}{dw}[t^{-1}(w)]\right| = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)}\left(1 + \frac{\left(\frac{w-\mu}{\sigma}\right)^2}{n}\right)^{-\frac{n+1}{2}}\frac{1}{\sigma} = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n \sigma^2}\,\Gamma\left(\frac{n}{2}\right)}\left(1 + \frac{1}{n}\frac{(w-\mu)^2}{\sigma^2}\right)^{-\frac{n+1}{2}}$$