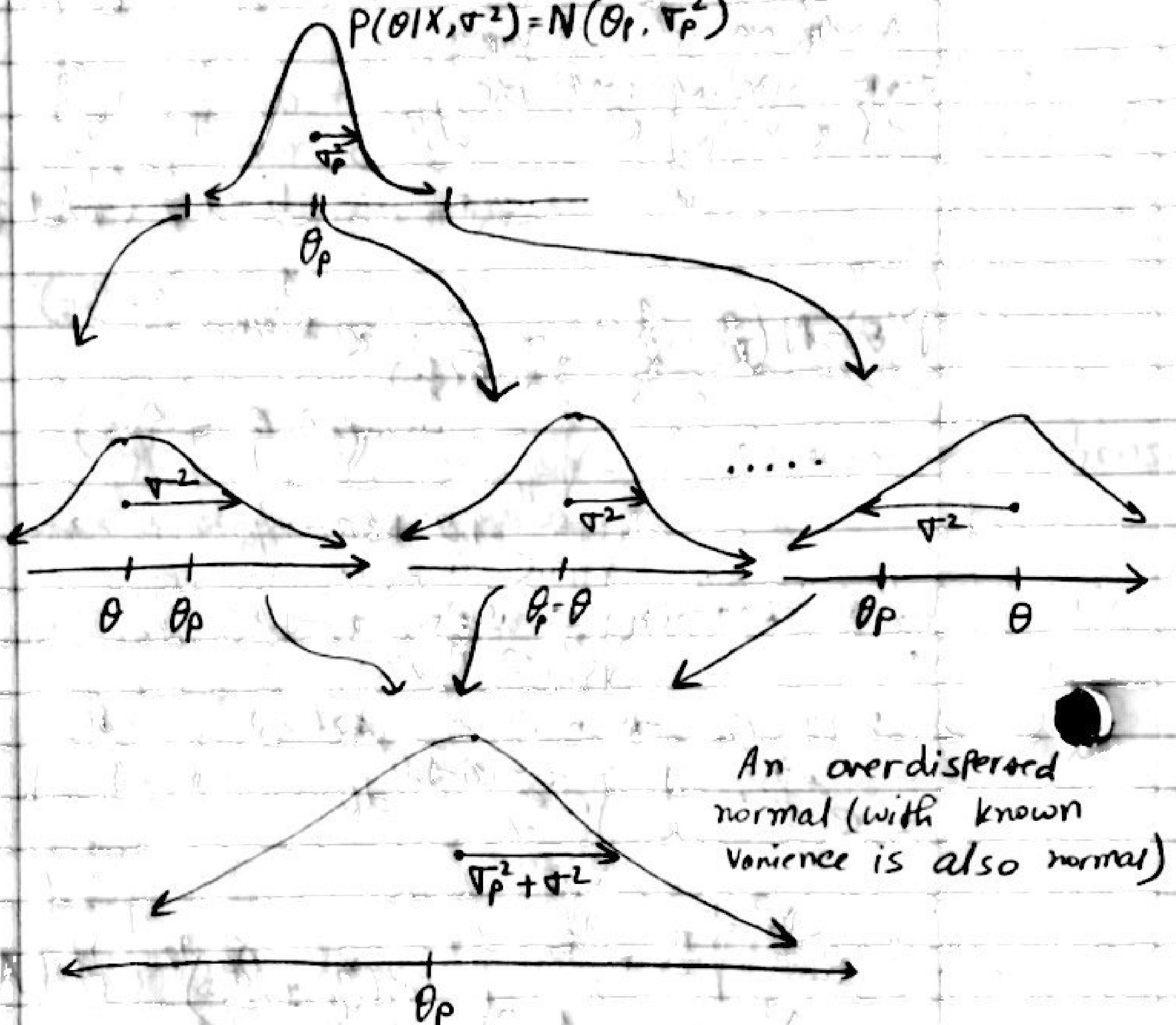
MATH 341/650.3

Lec 17

4/14/21



An overdispersed normal (with known variance is also normal)

Now consider the iid normal model with mean, θ, known, σ^2 unknown. i.e.

$\mathcal{F}: \overset{iid}{\sim} N(\theta, \sigma^2)$, θ known

$$P(X \mid \theta, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(X_i - \theta)^2}$$

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

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

$$\propto (\sigma^2)^{-n/2} e^{-\frac{\Sigma(X_i - \theta)^2/2}{\sigma^2}}$$

$$= (\sigma^2)^{\frac{-n}{2}} e^{\frac{-n\,\hat{\theta}^2_{MLE}/2}{\sigma^2}}$$

Consider the Laplace Prior of indiffrence, a distribution on σ^{-2} which has support $(0, \infty)$. This Prior would be... $P(\sigma^2|\theta) \propto 1$

Let's take a break and find the MLE for σ^{-2}

$$l(\sigma^2; x, \theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\Sigma(x_i-\theta)^2$$

$$l'(\sigma^2; x, \theta) = -\frac{n}{2\sigma^2} + \frac{\Sigma(x_i-\theta)^2}{2(\sigma^2)^2} \overset{set}{=} 0$$

$$\Rightarrow \frac{\Sigma(x_i-\theta)^2}{\sigma^2} = n \Rightarrow \hat{\sigma}^2_{MLE} = \frac{\Sigma(x_i-\theta)^2}{n}$$

Avg sqd derivation.

Let's explore the kernel of the posterior using probability theory.

$$k(y) = y^{-a} e^{-\frac{b}{y}}$$

Let's try to find the actual density by finding the norm. constant c:

$$\frac{1}{c} = \int_0^\infty k(y)dy = \int_0^\infty y^{-a} e^{-\frac{b}{y}} dy$$

let $z = \frac{1}{y} \Rightarrow y = \frac{1}{z} \Rightarrow \frac{dy}{dz} = z^{-2} \Rightarrow dy = -z^{-2}dz$

$y = 0 \Rightarrow z = \infty, y = \infty \Rightarrow z = 0$

$$\hookrightarrow = \int_\infty^0 z^a e^{-bz}(-z^{-2})dz = \int_0^\infty z^{a-1-1} e^{-bz} dz$$

$$\underset{=}{\text{u-sub}} \frac{\Gamma(a-1)}{b^{a-1}} \Rightarrow P(y) = \frac{b^{a-1}}{\Gamma(a-1)} y^a e^{-\frac{b}{y}}$$

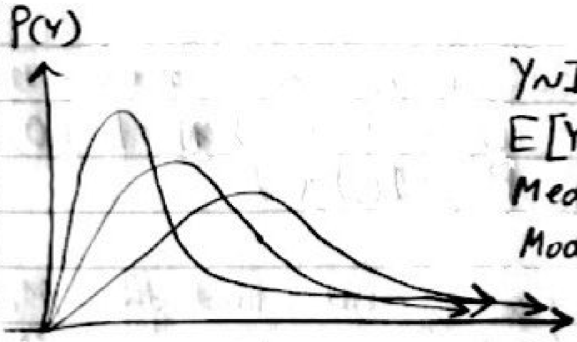
Traditionally, $\alpha = a-1$ $\Rightarrow a = \alpha+1$ $\beta = b$

$$\Rightarrow P(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-\alpha-1} e^{-\frac{\beta}{y}}$$

$$= \text{InvGamma}(\alpha, \beta) ; \alpha, \beta > 0$$

This is called the "inverse gamma" distribution.

Note. $W \sim \text{Gamma}(\alpha, \beta) \Leftrightarrow \frac{1}{W} \sim \text{Inv Gamma}(\alpha, \beta)$



$Y \sim InvGamma(\alpha, \beta)$
$E[Y] = \frac{\beta}{\alpha - 1}$ for $\alpha > 1$
$Med[Y] = $ qinvgamma$(0.5, \alpha, \beta)$
$Mode[Y] = \frac{\beta}{\alpha + 1}$ for all $\alpha, \beta > 0$

Back to the regular schedule program .....

$$P(\sigma^2|X,\theta) \propto (\sigma^2)^{-n/2} e^{-\frac{n\hat{\sigma}^2_{MLE}/2}{\sigma}} = (\sigma^2)^{-\frac{n-2}{2}-1} e^{-\frac{n\hat{\sigma}^2_{MLE}/2}{\sigma^2}}$$

$$\boxed{-\frac{n}{2} = -\frac{n}{2} + 1 - 1 = -\left(\frac{n}{2} - 1\right) - 1 = -\frac{n-2}{2} - 1} \propto InvGamma\left(\frac{n-2}{2}, \frac{n\hat{\sigma}^2_{MLE}}{2}\right)$$

That's the posterior under Laplace's Prior. Let's get the conjugate model now:

$$P(\sigma^2|X,\theta) \propto P(X|\theta,\sigma^2)P(\sigma^2|\theta) \propto (\sigma^2)^{-\frac{n}{2}} e^{\frac{-n\hat{\sigma}^2_{MLE}/2}{\sigma^2}} P(\sigma^2|\theta)$$

What form should the prior be so that it's kernel has the same form as the posterior's kernel? It's an inverse Gamma.

Let $P(\sigma^2|\theta) = InvGamma(\alpha, \beta)$

$$= (\sigma^2)^{-n/2} e^{-\frac{n\hat{\sigma}^2_{MLE}/2}{\sigma^2}} \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} e^{-\frac{\beta}{\sigma^2}}$$

$$\propto (\sigma^2)^{-\left(\frac{n}{2}+\alpha\right)-1} e^{-\frac{\frac{n\hat{\sigma}^2_{MLE}}{2}+\beta}{\sigma^2}}$$

$$\propto InvGamma\left(\frac{n}{2}+\alpha, \frac{n\hat{\sigma}^2_{MLE}}{2}+\beta\right)$$

Traditionally, we use a diffrent parameterization of the prior:

Let $\alpha = \frac{n_0}{2}$, $\beta = \frac{n_0\sigma_0^2}{2} \Rightarrow P(\sigma^2|\theta) = InvGamma\left(\frac{n_0}{2}, \frac{n_0\sigma_0^2}{2}\right)$

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

Bayesian Point estimates for σ^2:

$\alpha > 1 \Rightarrow \frac{n+n_0}{2} > 1 \Rightarrow$

$$\hat{\theta}_{MMSE} = E[\sigma^2|X,\theta] = \frac{\frac{n\hat{\theta}^2_{MLE}+n_0\sigma_0^2}{2}}{\frac{n+n_0}{2}-1} = \frac{n\hat{\sigma}^2_{MLE}+n_0\sigma_0^2}{n+n_0-2} \text{ if } n+n_0 > 2$$

$\hat{\theta}_{MMAE} = Med[\sigma^2 | X, \theta] = \text{qinvgamma}\left(0.5, \frac{n+n_0}{2}, \frac{n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}{2}\right)$

$\hat{\theta}_{MAP} = Mode[\sigma^2 | X, \theta] = \cdots = \frac{n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}{n+n_0+2}$

Credible regions? Same thing... Just use appropriate qinvgamma.

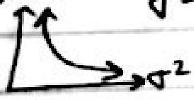
Hypothesis regions? Same thing... Just use appropriate pinvgamma.

Pseudoobservation interpretation n_0 = # of pseudo observation

Imagine $Y_1, Y_2 \ldots Y_{n_0}$. σ_0^2 = guess of value of σ^2 $\Rightarrow \sigma_0^2 = \frac{1}{n_0}\sum(X_i - \theta)^2$

$$\frac{n\hat{\sigma}^2_{MLE} + n_0\sigma_0^2}{2} = \frac{\sum_{i=1}^{n}(X_i-\theta)^2 + \sum_{i=1}^{n_0}(Y_i-\theta)^2}{2}$$

Haldane's Prior of absolute ignorance: $\frac{0^0}{\Gamma(0)}(\sigma^2)^{-0-1}e^{-\frac{0}{\sigma^2}} = \frac{1}{\sigma^2}$



$n_0 = 0 \Rightarrow P(\sigma^2|\theta) = \text{InvGamma}(0,0)$

σ_0^2 can be anything so by convention we say 0.

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

Laplace's Prior of indifference: $P(\sigma^2|\theta) \propto 1$

Is this a smart idea?
This means that σ^2 in $[0, 1]$ has the same weight as σ^2 in $[10^9, 1000000001]$.

This is not a smart idea. Noone really uses this prior.

What does this Laplace Prior correspond to? Recall it results in a Posterior of

$P(\sigma^2|X,\theta) = \text{InvGamma}\left(\frac{n-2}{2}, \frac{n\hat{\sigma}^2_{MLE}}{2}\right) \Rightarrow \begin{matrix} n_0 = -2 \\ \sigma^2 = 0 \end{matrix}$

$P(\sigma^2|\theta) = \text{InvGamma}\left(\frac{-2}{2}, \frac{0}{2}\right) = \text{InvGamma}(-1, 0)$

Scanned with CamScanner