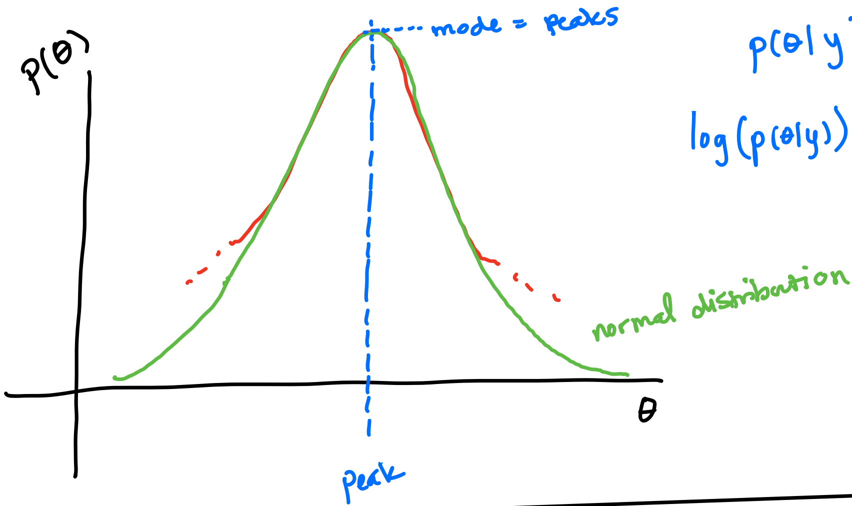

$$p(\theta|y) \propto \exp(a\theta^2 + b\theta + c)$$

$$\log(p(\theta|y)) \propto a\theta^2 + b\theta + c$$

$$\frac{d}{d\theta} \log[p(\theta|y)]\Big|_{\theta=\hat{\theta}} = 0$$

θ is likely to be multidimensional

$$f(\theta) \approx f(\theta_0) + \nabla f(\theta_0)^T(\theta - \theta_0) + \frac{1}{2}(\theta - \theta_0)^T H f(\theta_0)(\theta - \theta_0)$$

$$\log(p(\theta|y)) \approx \log p(\hat{\theta}|y) + 0 + \frac{1}{2}(\theta - \hat{\theta})^T \left[\frac{d^2}{(d\theta)^2} \log p(\theta|y)\right]_{\theta=\hat{\theta}} (\theta - \hat{\theta})$$

observed Fisher information

$$I(\theta) = -\frac{\partial^2}{\partial\theta^2} \log p(\theta|y)$$

$$\log(p(\theta|y)) \approx \log p(\hat{\theta}|y) + 0 + \frac{1}{2}(\theta-\hat{\theta})^T \left[\frac{d^2}{(d\theta)^2}\log p(\theta|y)\right]_{\theta=\hat{\theta}}(\theta-\hat{\theta})$$

$$\log p(\theta|y) \approx A + \frac{1}{2}(\theta-\hat{\theta})^T I(\theta)(\theta-\hat{\theta})$$

$$p(\theta|y) \approx B e^{\frac{1}{2}(\theta-\hat{\theta})^T I(\theta)(\theta-\hat{\theta})}$$

multiparameter normal distribution

$$p(\theta|y) \approx N(\hat{\theta}, I(\theta)^{-1})$$

<u>Example</u>: normal distribution with unknown and unknown variance

<u>Prior</u>: noninformative prior $p(\mu,\sigma^2) \propto \frac{1}{\sigma^2}$

show $p(\mu, \log\sigma) \propto 1$

$(\mu,\sigma^2) \mapsto (\mu, \log\sigma)$

$$J = \begin{pmatrix} \frac{\partial\mu}{\partial\mu} & \frac{\partial\mu}{\partial\log\sigma} \\ \frac{\partial\sigma^2}{\partial\mu} & \frac{\partial\sigma^2}{\partial\log\sigma} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2\sigma^2 \end{pmatrix}$$

$\hookrightarrow \sigma^2 = e^{2\log\sigma}$

$$\frac{\partial e^{2\log\sigma}}{\partial\log\sigma} = 2e^{2\log\sigma} = 2\sigma^2$$

$|J| = 2\sigma^2$ So $p(\mu,\log\sigma) = p(\mu,\sigma^2)\cdot|J| \propto \frac{1}{\sigma^2}\cdot 2\sigma^2$

$= 2 \propto 1$

① $\theta = (\mu, \log\sigma)$ so our noninformative prior is

$p(\mu, \log\sigma) = 1$

② Find $\hat{\theta}$.

Need to find the peak of $p(\mu, \log\sigma \mid y)$

Recall from module 3:

due to [J]

$$p(\mu, \cancel{\sigma^2}^{\log\sigma} \mid y) \propto \sigma^{-n\cancel{-2}} \exp\left[-\frac{1}{2\sigma^2}\left((n-1)s^2 + n(\bar{y}-\mu)^2\right)\right]$$

$$\boxed{\log(p(\mu, \log\sigma \mid y)) \propto -n\log\sigma - \frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]}$$

this is the function around which we will construct our Taylor series

$\hat{\theta}$ is the point where $\nabla \log(p(\mu, \log\sigma \mid y)) = \vec{0}$

$$\frac{\partial}{\partial\mu} \log(p(\mu, \log\sigma \mid y)) \propto \frac{n(\bar{y}-\mu)}{\sigma^2} = 0$$

$\sigma^{-2} = e^{-2\log\sigma}$

$$\frac{\partial}{\partial\log\sigma} \log(p(\mu, \log\sigma \mid y)) \propto -n + \frac{(n-1)s^2 + n(\bar{y}-\mu)^2}{\sigma^2} = 0$$

$$\frac{n(\bar{y}-\mu)}{\sigma^2} = 0 \Rightarrow n(\bar{y}-\mu) = 0 \Rightarrow \boxed{\bar{y} = \mu} = \hat{\theta}_1$$

$$-n + \frac{(n-1)s^2 + n(\cancel{\bar{y}-\mu})^2}{\sigma^2} = 0 \quad \Rightarrow \quad \frac{-n\sigma^2 + (n-1)s^2}{\sigma^2} = 0$$

$$\Rightarrow (n-1)s^2 = n\sigma^2$$

$$\Rightarrow \boxed{\frac{(n-1)s^2}{n} = \sigma^2} \neq \hat{\theta}_2$$

$$2\log\sigma = \log\left(\frac{(n-1)s^2}{n}\right) \quad \Rightarrow \quad \boxed{\log\sigma = \frac{1}{2}\log\left(\frac{(n-1)s^2}{n}\right)} = \hat{\theta}_2$$

$$\boxed{\hat{\theta} = \begin{pmatrix} \bar{y} \\ \frac{1}{2}\log\left(\frac{(n-1)s^2}{n}\right) \end{pmatrix}}$$

③ Find $I(\hat{\theta})$.

$$\boxed{\log(p(\mu, \log\sigma \mid y)) \propto -n\log\sigma - \frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]}$$

$$\frac{\partial}{\partial \mu} \log\left(p(\mu, \log\sigma \mid y)\right) \propto \frac{n(\bar{y}-\mu)}{\sigma^2}$$

$$\frac{\partial}{\partial \log\sigma} \log\left(p(\mu, \log\sigma \mid y)\right) \propto -n + \frac{(n-1)s^2 + n(\bar{y}-\mu)^2}{\sigma^2}$$

$$\frac{\partial^2}{(\partial \mu)^2} \log\left(p(\mu, \log\sigma \mid y)\right) = \frac{-n}{\sigma^2}$$

$@\ \hat{\theta}: \ -n \cdot \frac{n}{(n-1)s^2} = \frac{-n^2}{(n-1)s^2}$

$$\frac{\partial^2}{\partial \mu \, \partial \log\sigma} \log\left(p(\mu, \log\sigma \mid y)\right) = -2n(\bar{y}-\mu)e^{-2\log\sigma} = \frac{-2n(\bar{y}-\mu)}{\sigma^2}$$

$@\ \hat{\theta}: 0$

$$\frac{\partial^2}{(\partial \log\sigma)^2} \log\left(p(\mu, \log\sigma \mid y)\right) = -2\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]e^{-2\log\sigma}$$

$@\ \hat{\theta}: \ \frac{-2(n-1)s^2}{\sigma^2} = \frac{-2ns^2}{s^2} = -2n$

$$\boxed{-I(\hat{\theta}) = \begin{pmatrix} -\frac{n^2}{(n-1)s^2} & 0 \\ 0 & -2n \end{pmatrix}}$$

(f) Compute the normal distribution approximation

$$p(\theta \mid y) \approx N\left(\hat{\theta}, I(\hat{\theta})^{-1}\right)$$

$$p(\mu, \log\sigma \mid y) \approx N\left(\begin{pmatrix} \bar{y} \\ \frac{1}{2}\log\left(\frac{(n-1)s^2}{2}\right) \end{pmatrix}, \begin{pmatrix} \frac{n^2}{(n-1)s^2} & 0 \\ 0 & 2n \end{pmatrix}^{-1} \right)$$

Solution to the example