Xinyu Tan

– **1. Normal Approximations**

(a) Log posterior density:

$$\begin{aligned}\log p(\theta|y_1,\cdots,y_5) &= \log p(\theta)\log(y1,\cdots,y_5|\theta)\\ &= \sum_{i=1}^{5}\log p(y_i|\theta) \propto \sum_{i=1}^{5}\log\frac{1}{1+(y_i-\theta)^2}\\ &= -\sum_{i=1}^{5}\log\left(1+(y_i-\theta)^2\right)\end{aligned}$$

Hence, first derivative:

$$\frac{dp(\theta|y)}{d\theta} = 2\sum_{i=1}^{5}\frac{y_i-\theta}{1+(y_i-\theta)^2}$$

Second derivative:

$$\frac{d^2p(\theta|y)}{d\theta^2} = 2\sum_{i=1}^{5}\frac{(y_i-\theta)^2-1}{(y_i-\theta)^2+1}$$

(b) The posterior mode $\hat{\theta} = -0.125$

(c) The posterior normal approximation:

$$\begin{aligned}\log p(\theta|y) &\approx p(\hat{\theta}|y) + \frac{1}{2}(\theta-\hat{\theta})^2 \times \left[\frac{d^2p(\theta|y)}{d\theta^2}\right]_{\theta=\hat{\theta}}\\ &= -5.45 + \frac{1}{2}\times 1.30 \times (\theta+0.125)^2\end{aligned}$$

Therefore, the approximated posterior distribution is $\theta|y \sim N(-0.125, 0.877^2)$.

– **2. Normal Approximation**

Note: *Trivial arithmetic*

In bioassay example, we have the posterior

$$p(\alpha,\beta|y) = \prod_{i=1}^{4}\textbf{logit}^{-1}(\alpha+\beta x_i)^{y_i}\left(1-\textbf{logit}^{-1}(\alpha+\beta x_i)\right)^{n_i-y_i}$$

Denote logit$^{-1}(x)$ to be $f(x)$; hence, $f(x) = 1/\left(1+e^{-x}\right)$, the derivative $df(x)/dx = (1-f)f$. Using $f(x)$, the log likelihood:

$$\log p(\alpha,\beta|y) = \sum_{i=1}^{4} y_i\log f(\alpha+\beta x_i) + (n_i-y_i)\log\left(1-f(\alpha+\beta x_i)\right)$$

First compute mode:

$$\begin{aligned}\frac{\partial\log p(\alpha,\beta|y)}{\partial\alpha} &= \sum_{i=1}^{4} y_i - n_i f(\alpha+\beta x_i)\\ \frac{\partial\log p(\alpha,\beta|y)}{\partial\beta} &= \sum_{i=1}^{4} x_i(y_i - n_i f(\alpha+\beta x_i))\end{aligned}$$

Second derivative:

$$\frac{\partial^2 \log p(\alpha,\beta|y)}{\partial\alpha^2} = -\sum_{i=1}^{4} n_i f(\alpha+\beta x_i)\left(1-f(\alpha+\beta x_i)\right) \tag{1}$$

$$\frac{\partial^2 \log p(\alpha,\beta|y)}{\partial\alpha\partial\beta} = -\sum_{i=1}^{4} n_i x_i f(\alpha+\beta x_i)(1-f(\alpha+\beta x_i)) \tag{2}$$

$$\frac{\partial^2 \log p(\alpha,\beta|y)}{\partial\alpha\partial\beta} = -\sum_{i=1}^{4} n_i x_i^2 f(\alpha+\beta x_i)(1-f(\alpha+\beta x_i)) \tag{3}$$

The information matrix

$$I(\alpha,\beta) = \begin{bmatrix} \sum_{i=1}^4 n_i f(\alpha+\beta x_i)\left(1-f(\alpha+\beta x_i)\right) & \sum_{i=1}^4 n_i x_i f(\alpha+\beta x_i)(1-f(\alpha+\beta x_i)) \\ \sum_{i=1}^4 n_i x_i f(\alpha+\beta x_i)(1-f(\alpha+\beta x_i)) & \sum_{i=1}^4 n_i x_i^2 f(\alpha+\beta x_i)(1-f(\alpha+\beta x_i)) \end{bmatrix}$$

– **9. Point estimation**

I did not figure this problem out. I looked at the answer key. And then the following is my understanding from the answer. This is a really nice problem - linking many concepts together.

Suppose the measurement is $\hat{y}$. The likelihood

$$\phi(\theta) = p(\hat{y}|\theta) = \frac{1}{\sqrt{2\pi}\sigma}\mathbf{exp}\left(-\frac{1}{2\sigma^2}(\hat{y}-\theta)^2\right)$$

here is a part I did not think of. So basically, it should be assumed that when $\theta < 0$, with $\theta \in [0,1]$, then $p(\theta < 0)$ should be assigned to $p(\theta = 0)$.

Then,

$$\begin{aligned}\phi(\theta=0) &= \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}\sigma}\mathbf{exp}\left(-\frac{1}{2\sigma^2}(\hat{y}-\theta)^2\right)d\theta \\ &= \frac{1}{2}\left(1+\mathbf{erf}\left(\frac{-\hat{y}}{\sigma\sqrt{2}}\right)\right)\end{aligned}$$

and

$$\begin{aligned}\phi(\theta=1) &= \int_{1}^{\infty} \frac{1}{\sqrt{2\pi}\sigma}\mathbf{exp}\left(-\frac{1}{2\sigma^2}(\hat{y}-\theta)^2\right)d\theta \\ &= 1-\frac{1}{2}\left(1+\mathbf{erf}\left(\frac{1-\hat{y}}{\sigma\sqrt{2}}\right)\right)\end{aligned}$$

$\mathbf{erf}(0) = 0$, then when $\sigma \to \infty$, $\phi(0) \to 1/2$ and $\phi(1) \to 1/2$.
Therefore, in the case of max likelihood, we have θ's estimate to be either 1 or 0:

$$\mathbf{MSE}_1 = 0.5 \times \theta^2 + 0.5 \times (1-\theta)^2 = \frac{1}{2}(1-2\theta+2\theta^2)$$

Posterior distribution:

$$\psi(\theta) = p(\theta|\hat{y}) \propto I(0 \le \theta \le 1)\frac{1}{\sqrt{2\pi}\sigma}\mathbf{exp}\left(-\frac{1}{2\sigma^2}(\hat{y}-\theta)^2\right)$$

The posterior mean

$$\begin{aligned}\bar{\theta} &= \frac{\int_0^1 \theta\psi(\theta)d\theta}{\int_0^1 \psi(\theta)d\theta} \\ &\approx \int_0^1 \theta d\theta \qquad (\sigma \to \infty) \\ &= \frac{1}{2}\end{aligned}$$

Hence,

$$\mathbf{MLE}_2 = (0.5-\theta)^2 = \frac{1}{2}\left(\frac{1}{2} - 2\theta + 2\theta^2\right) < \mathbf{MLE}_1$$

where $\theta \in [0, 1]$