Reliability

It is good practice to report a point estimate (single number/vector summary such as a posterior mean or posterior mode) together with a measure of reliability (CI, HPD region, posterior variance, Laplace approx.

Confidence Intervals

Frequentist CI is a probability statement about the interval. i.e. $p(\ell(Y) < \theta < u(Y) \mid \theta) = (1-\alpha)$

$\uparrow$ random $\qquad \uparrow$ random

Bayesian CI is a probability statement about θ.

i.e. $p(\ell(y) < \theta < u(y) \mid Y=y) = (1-\alpha)$

$\uparrow$ observed data

Practical way to compute Bayesian CI @ $(1-\alpha)$ level.

- compute or approximate posterior
- find quantiles $\ell(y)$ and $u(y)$ such that

$p(\theta < \ell(y)) = \frac{\alpha}{2}$

$p(\theta < u(y)) = 1 - \frac{\alpha}{2}$

so that $p(\theta \in (\ell(y), u(y))) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha$.

Ex: beta-binomial posterior

likelihood : $p(y_1, \dots, y_n \mid \theta) \propto \theta^{\sum y_i} (1-\theta)^{n - \sum y_i}$

prior : $p(\theta) \propto \theta^{a-1} (1-\theta)^{b-1}$

posterior : $p(\theta \mid y_1, \dots, y_n) \propto \theta^{\alpha - 1} (1-\theta)^{\beta - 1}$

where $\alpha = a + \sum y_i$

$\beta = n + b - \sum y_i$

95% posterior CI:

qbeta(c(0.025, 0.975), α, β)

Aside: posteriors combine prior info. w/ data info.

ex: posterior mean of beta-binomial

$\theta \mid y_1, \ldots, y_n \sim \text{beta}(\alpha, \beta)$

$$\mathbb{E}\,\theta \mid y_1, \ldots, y_n = \frac{\alpha}{\alpha+\beta} = \frac{a+\Sigma y_i}{a+b+n}$$

Recall: $\bar{y} = \frac{1}{n}\Sigma y_i \Rightarrow n\bar{y} = \Sigma y_i$

$$\mathbb{E}\,\theta \mid y_1, \ldots, y_n = \frac{a+n\bar{y}}{a+b+n}$$

$$= \frac{a}{a+b+n} \cdot \frac{(a+b)}{(a+b)} + \frac{n}{a+b+n}\bar{y}$$

$$= \underbrace{\boxed{\frac{a+b}{a+b+n}}}_{\substack{\text{weight} \\ w}} \underbrace{\boxed{\frac{a}{(a+b)}}}_{\substack{\text{prior mean} \\ \mathbb{E}\theta}} + \underbrace{\boxed{\frac{n}{a+b+n}}}_{(1-w)} \underbrace{\boxed{\bar{y}}}_{\text{sample mean}}$$

Laplace Approximation:

Fit a normal (Gaussian) to the mode of the posterior

Method: Taylor expand $\underbrace{\log p(\theta \mid \vec{y})}_{\text{log-posterior}}$ about $\hat{\theta}_{MAP}$.

Where $\hat{\theta}_{MAP} = \underset{\theta}{\operatorname{argmax}} \; p(\theta \mid \vec{y})$

i.e. $\hat{\theta}_{MAP}$ s.t. $\frac{d}{d\theta} \log p(\theta \mid \vec{y}) = 0$ and

locally concave, $\frac{d^2}{d\theta^2} \log p(\theta \mid \vec{y}) \Big|_{\hat{\theta} \pm \epsilon} < 0$

<u>Define</u>

Let $\log p(\theta \mid \vec{y}) = L(\theta)$ for convenience.

Let $\hat{\theta}_{MAP} := \hat{\theta}$ for convenience.

$$L(\theta) \approx L(\hat{\theta}) + \underbrace{\cancel{L'(\hat{\theta})}}_{0}(\theta - \hat{\theta}) + \frac{1}{2} L''(\hat{\theta})(\theta - \hat{\theta})^2$$

$$p(\theta \mid \vec{y}) = e^{L(\theta)} \approx \underbrace{e^{L(\hat{\theta})} e^{\frac{1}{2} L''(\hat{\theta})(\theta - \hat{\theta})^2}}_{\text{kernel of normal}}$$

w/ mean $= \hat{\theta}$

variance $= -1/L''(\hat{\theta})$

So $\theta \mid \vec{y} \approx N\left(\hat{\theta}_{map}, -1/L''(\hat{\theta}_{map})\right)$

Laplace Approx Example:

Approximate beta-binomial posterior

$$p(\theta \mid \vec{y}) = c\,\theta^{\alpha-1}(1-\theta)^{\beta-1}$$

$$L(\theta) = \log \theta \mid \vec{y} = \log c + (\alpha-1)\log\theta + (\beta-1)\log(1-\theta)$$

$$L'(\theta) = \frac{\alpha-1}{\theta} - \frac{(\beta-1)}{(1-\theta)}$$

$$\boxed{L''(\theta) = \frac{-(\alpha-1)}{\theta^2} - \frac{(\beta-1)}{(1-\theta)^2}}$$

Set $L'(\theta) = 0$ to find $\boxed{\hat{\theta}_{map} = \dfrac{\alpha-1}{\beta+\alpha-2}}$

So

$$p(\theta \mid y) \approx N\left(\hat{\theta}_{map},\ {-1}/{L''(\hat{\theta}_{map})}\right)$$

where the relevant terms $\hat{\theta}_{map}$ & $L''(\theta)$ are circled above.