ADVANCED BAYESIAN MODELING

DIC for 2016 Polls Data: Explicit Cases

2016 Presidential Polls Example

Seven polls conducted within days of 2016 U.S. Presidential election for two-way race (Clinton v. Trump):

$$\begin{aligned} y_j &= \text{Clinton lead (percentage points) in poll } j \\ \sigma_j &= \text{half margin of error of } y_j \\ j &= 1, \ldots, 7 \end{aligned}$$

Regard σ_js as fixed and known.

```
> d <- read.table("polls2016.txt", header=TRUE)

> d$sigma <- d$ME/2  # standard dev = half margin of error

> d
       poll y   ME sigma
1    YouGov 4 1.70 0.850
2 Bloomberg 3 3.50 1.750
3   ABCWaPo 3 2.50 1.250
4       Fox 4 2.50 1.250
5       IBD 1 3.10 1.550
6  Monmouth 6 3.60 1.800
7    NBCWSJ 5 2.73 1.365
```

Recall hierarchical model:

$$\begin{aligned}
y_j \mid \theta_j &\sim \mathrm{N}(\theta_j, \sigma_j^2) \qquad j = 1, \ldots, 7 \\
\theta_j \mid \mu, \tau &\sim \mathrm{N}(\mu, \tau^2) \qquad j = 1, \ldots, 7 \\
\mu &\sim \text{flat on } (-\infty, \infty) \\
\tau &\sim \text{flat on } (0, \infty)
\end{aligned}$$

Goal: "Pool" information about θs from different polls to get better estimates and (perhaps) predictions.

Two alternative models of interest:

- No Pooling ($\tau \to \infty$):

  Joint prior on $(\theta_1, \ldots, \theta_7)$ is flat.

  As if polls are analyzed independently (each with a flat mean prior)

- Complete Pooling ($\tau = 0$):

  All θ_js equal a single common mean μ having a flat prior.

  All polls assumed to be estimating the same Clinton lead.

No Pooling Model

$$y_j \mid \theta_j \sim \mathrm{N}(\theta_j, \sigma_j^2) \qquad j = 1, \ldots, 7$$

$$\theta_j \sim \text{flat on } (-\infty, \infty) \qquad j = 1, \ldots, 7$$

These are completely independent mean-only single-observation normal models with flat priors, so posteriors are

$$\theta_j \mid y \sim \text{indep. } \mathrm{N}(y_j, \sigma_j^2) \qquad j = 1, \ldots, 7$$

The deviance:

$$
\begin{aligned}
-2\,\log p(y \mid \theta) &= -2\,\log \prod_{j=1}^{7} p(y_j \mid \theta) \;=\; -2 \sum_{j=1}^{7} \log p(y_j \mid \theta_j) \\
&= -2 \sum_{j=1}^{7} \left(-\log \sqrt{2\pi\sigma_j^2} - \frac{1}{2\sigma_j^2}(y_j - \theta_j)^2 \right) \\
&= \sum_{j=1}^{7} \log\bigl(2\pi\sigma_j^2\bigr) + \sum_{j=1}^{7} \frac{(y_j - \theta_j)^2}{\sigma_j^2}
\end{aligned}
$$

Evaluating deviance at $\hat{\theta}_{\mathsf{Bayes}} = \mathrm{E}(\theta \mid y) = y$ gives

$$
\begin{aligned}
-2 \log p\big(y \mid \hat{\theta}_{\mathsf{Bayes}}\big) &= \sum_{j=1}^{7} \log\big(2\pi\sigma_j^2\big) + \sum_{j=1}^{7} \frac{(y_j - y_j)^2}{\sigma_j^2} \\
&= \sum_{j=1}^{7} \log\big(2\pi\sigma_j^2\big) \approx 17.2
\end{aligned}
$$

Averaging deviance over the posterior gives

$$
\begin{aligned}
\mathrm{E}_{\mathsf{post}}\big(-2\,\log p(y \mid \theta)\big) &= \sum_{j=1}^{7} \log\big(2\pi\sigma_j^2\big) \;+\; \sum_{j=1}^{7} \frac{\mathrm{E}_{\mathsf{post}}(\theta_j - y_j)^2}{\sigma_j^2} \\
&= \sum_{j=1}^{7} \log\big(2\pi\sigma_j^2\big) \;+\; \sum_{j=1}^{7} \frac{\mathrm{var}_{\mathsf{post}}(\theta_j)}{\sigma_j^2} \\
&= \sum_{j=1}^{7} \log\big(2\pi\sigma_j^2\big) \;+\; \sum_{j=1}^{7} \frac{\sigma_j^2}{\sigma_j^2} \\
&= \sum_{j=1}^{7} \log\big(2\pi\sigma_j^2\big) \;+\; 7
\end{aligned}
$$

Thus, effective number of parameters is

$$\begin{aligned}
p_{\mathrm{DIC}} &= 2\left(\log p\big(y \mid \hat{\theta}_{\mathsf{Bayes}}\big) \;-\; \mathrm{E}_{\mathsf{post}}\big(\log p(y \mid \theta)\big)\right) \\
&= -\Big(-2\,\log p\big(y \mid \hat{\theta}_{\mathsf{Bayes}}\big)\Big) \;+\; \mathrm{E}_{\mathsf{post}}\big(-2\,\log p(y \mid \theta)\big) \\
&= -\sum_{j=1}^{7} \log\big(2\pi\sigma_j^2\big) \;+\; \sum_{j=1}^{7} \log\big(2\pi\sigma_j^2\big) \;+\; 7 \;=\; 7
\end{aligned}$$

and

$$\mathrm{DIC} \;=\; -2\,\log p\big(y \mid \hat{\theta}_{\mathsf{Bayes}}\big) \;+\; 2p_{\mathrm{DIC}} \;\approx\; 17.2 \;+\; 2 \cdot 7 \;=\; 31.2$$

Remarks:

- Can show

$$p_{\mathrm{DICalt}} \;=\; 2\ \mathrm{var}_{\textsf{post}}\big(\log p(y \mid \theta)\big) \;=\; 7$$

  so the alternative DIC has the same value in this case.

- Can show AIC is the same as DIC in this case.

Complete Pooling Model

$$
\begin{aligned}
y_j \mid \mu \quad &\sim \quad \mathrm{N}(\mu, \sigma_j^2) \qquad\qquad j = 1, \ldots, 7 \\
\mu \quad &\sim \quad \text{flat on } (-\infty, \infty)
\end{aligned}
$$

This is a mean-only unequal-variance normal model (alternatively, intercept-only weighted normal-theory regression model) with flat prior.

Can be shown that

$$-2\,\log p\big(y \mid \hat{\mu}_{\textsf{Bayes}}\big) \;=\; \sum_{j=1}^{7} \log\big(2\pi\sigma_j^2\big) \;+\; \sum_{j=1}^{7} \frac{(y_j - \hat{\mu}_{\textsf{Bayes}})^2}{\sigma_j^2} \;\approx\; 23.4$$

where

$$\hat{\mu}_{\textsf{Bayes}} \;=\; \mathrm{E}(\mu \mid y) \;=\; \frac{\sum_{j=1}^{7} y_j/\sigma_j^2}{\sum_{j=1}^{7} 1/\sigma_j^2} \;\approx\; 3.75$$

and that

$$p_{\mathrm{DIC}} \;=\; p_{\mathrm{DIC}\textsf{alt}} \;=\; 1$$

Then, for complete pooling,

$$\mathrm{DIC} \;=\; -2\,\log p\big(y \mid \hat{\mu}_{\textsf{Bayes}}\big) \;+\; 2p_{\mathrm{DIC}} \;\approx\; 23.4 \;+\; 2\cdot 1 \;=\; 25.4$$

(AIC is the same.)