ADVANCED BAYESIAN MODELING

A NORMAL HIERARCHICAL MODEL:

HIERARCHICAL MODEL FOR 2016 POLLS

Review

Observe:

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

Let

$$
y = (y_1, \ldots, y_7) \qquad \sigma = (\sigma_1, \ldots, \sigma_7)
$$

Regard σ_j as (estimate of) $\sqrt{\mathrm{var}(y_j)}$ in sampling distribution of y_j.

Model for Poll Results

Results come from separately conducted polls:

$$y_1, \ldots, y_7 \quad \text{independent (in sampling distribution)}$$

To a close approximation, they should be normally distributed:

$$y_j \mid \theta_j \quad \sim \quad \mathrm{N}(\theta_j, \text{?}\,)$$

Note: Poll j is allowed its own mean θ_j.

Since σ_j represents standard error of y_j,

$$y_j \mid \theta_j \sim \mathrm{N}(\theta_j, \sigma_j^2)$$

For simplicity, regard σ_j as fixed and known. (No need to include it in conditioning notation.)

(More elaborate analysis might put independent priors on σ_js.)

Model for Means

No reason to assume differences among polls (prior to seeing data).

Therefore, model poll means $\theta_1, \ldots, \theta_7$ as exchangeable.

Convenient to model them conditionally independent from a normal distribution:

$$\theta_j \mid \mu, \tau \quad \sim \quad \mathrm{N}(\mu, \tau^2)$$

Reasonable if no skew in population and no outliers.

(Also, recall normal prior is conjugate for normal sample mean.)

Hyperprior

Noninformative hyperprior proposed in BDA3, Sec. 5.4, for similar situation:

$$p(\mu, \tau) \propto 1 \qquad -\infty < \mu < \infty \quad \tau > 0$$

Obtained from multiplying flat priors for μ and τ:

$$\begin{aligned} p(\mu) &\propto 1 \qquad -\infty < \mu < \infty \\ p(\tau) &\propto 1 \qquad \tau > 0 \end{aligned}$$

Note: Improper, so requires checking that posterior is proper.

Warning: Use improper hyperpriors carefully!

For example, using

$$p(\log \tau) \quad \propto \quad 1$$

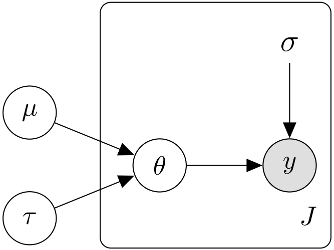
would give an improper posterior (BDA3, Sec. 5.4).

Full Model

The full hierarchical normal model:

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

DAG Model



Note: The **constant** node σ isn't circled.

Constant nodes are always observed.