# Hierarchical Model:

Motivation: $\theta_j$ = survival probability at hospital $j$ for a disease

observe: out of the $n_j$ patients admitted to the hospital $j$.

$y_j$ of the patients survived.

① for each $j$, $\longrightarrow$ single parameter model (Beta-Binomial model)

independently. ✓

② Assume $\theta_1 = \theta_2 = \theta_3 = \cdots$, $\Sigma n_j$ patients, $\Sigma y_j$ survived

(neglect different hospital)

$\longrightarrow$ separate model for each hospital

$\longrightarrow$ pooled model

in-between: hierarchical model

specify enough params to fit data well but not overfitting

---

Example: Rat Tumor

Quantity of interest:

$\theta$ = Prob ( tumor in female laboratory rats of type 'F344' that receives a 0 dose of the drug )

For a data, get to know which quantity we can / can't estimate!

observation: $\overset{y_{71}}{4}$ out of $\overset{n_{71}}{14}$ rats develop tumor $\rightarrow f(a,b|y_{71})$

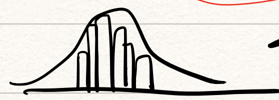
What other pieces of information is available?

— "historical data" $\{y_j, n_j\}_{j=1,2,\dots,70}$ laboratory

prior $\theta \sim Beta(\alpha, \beta)$ ← moments of $\{y_j/n_j\}$

↓ Beta - Binomial

posterior $\theta \sim Beta(\hat{\alpha}+4, \hat{\beta}+10)$



→ fit a Beta distribution and find $\hat{\alpha}, \hat{\beta}$

What if we also want to know each $\theta_j$ for $j=1,2,\dots,70$?

why × use $\theta_j \sim Beta(\hat{\alpha}+y_j, \hat{\beta}+n_j-y_j)$?

∵ we used this info in prior!

× use the data twice

And we're treating $\hat{\alpha}, \hat{\beta}$ as the truth! → wrong!

$(\alpha, \beta)$ global parameter / hyperparam

individual param.

$$
\begin{array}{cccc}
\theta_1 & \theta_2 & \dots & \theta_{70} & \theta_{71} \\
\downarrow & \downarrow & & \downarrow & \downarrow \\
y_1 & y_2 & & y_{70} & y_{71} \\
n_1 & n_2 & & n_{70} & n_{71}
\end{array}
$$

Model: for $j=1,2,\dots,70,71$

$$
\begin{cases}
\theta_j \sim Beta(\alpha, \beta) \\
n_j \sim Binomial(n_j, \theta_j)
\end{cases}
$$

$\boxed{\text{Key!}}$ posterior $\theta_j \mid \alpha, \beta, n_j, y_j \sim Beta(\alpha+y_j, \beta+n_j-y_j)$

Next. How to make inference for $(\alpha, \beta)$?

param of interest. nuisance → multi parameter model

$$P(\boxed{\alpha, \beta}, \boxed{\theta_{1:71}} \mid y_1, y_2, \cdots, y_{71}, n_1, n_2, \cdots, n_{71})$$

$$\propto P(\alpha, \beta) \boxed{\underbrace{P(\theta_{1:71} \mid \alpha, \beta)}_{\text{prior}} \underbrace{P(y_{1:71}, n_{1:71} \mid \theta_{1:71})}_{\text{likelihood}}}$$

Conditional

$(\alpha, \beta) \to \theta \to y, n$

Conditional independence

$$= P(\alpha, \beta) \prod_{j=1}^{71} \left[ \theta_j^{\alpha + y_j - 1} (1 - \theta_j)^{\beta + n_j - y_j - 1} \right]$$

Use the idea of multi-param model:

$\boxed{\text{Key2}}$

$$P(\alpha, \beta \mid y_{1:71}, n_{1:71}) = \int P(\alpha, \beta, \theta_{1:71} \mid y_{1:71}, n_{1:71}) \, d\theta_{1:71}$$

$$\propto P(\alpha, \beta) \prod_{j=1}^{71} \int \theta_j^{\alpha + y_j - 1} (1 - \theta_j)^{\beta + n_j - y_j - 1} \, d\theta_j$$

$$\propto P(\alpha, \beta) \prod_{j=1}^{71} \text{Beta}(\alpha + y_j, \beta + n_j - y_j)$$

It's easy from Bayes formula to derive $P(\alpha, \beta, \theta_{1:71} \mid y_{1:71}, n_{1:71})$ Not using data twice

73 dimensional posterior distribution

$P(\alpha, \beta \mid y_{1:71}, n_{1:71})$ : marginal posterior for global params

↳ uncertainty of $\alpha, \beta$

$P(\theta_j \mid \alpha, \beta, y_{1:71}, n_{1:71})$ : conditional posterior for individual params, conditioning on global params

Recall what we did in multi-model:

$$p(\alpha, \beta \mid y_{1:71}, n_{1:71}) \longrightarrow \text{grid approx/ MCMC}$$

$$\hookrightarrow \{\alpha^{(i)}, \beta^{(i)}\}_{i=1,2,\cdots,M}$$

for each $\{\alpha^{(i)}, \beta^{(i)}\}$, draw $\theta_j^{(i)} \sim \text{Beta}(\alpha^{(i)} + y_j, \beta^{(i)} + n_j - y_j)$

$$\hookrightarrow \{\theta_1^{(i)}\}_{i=1\cdots M}, \{\theta_2^{(i)}\}_{i=1\cdots M}, \cdots \{\theta_{71}^{(i)}\}_{i=1\cdots M}$$

posterior inference:

$$\{\alpha^{(i)}, \beta^{(i)}, \theta_1^{(i)}, \theta_2^{(i)}, \cdots \theta_{71}^{(i)}\}_{i=1\cdots M}$$

$$\hookrightarrow p(\alpha, \beta, \theta_{1:71} \mid y_{1:71}, n_{1:71})$$

$$\boxed{E[\theta_{71} \mid -] = \int \theta_{71}\, p(\alpha, \beta, \theta_{1:71} \mid -)\, d\alpha\, d\beta\, d\theta_{1:71}}$$

Summary