ADVANCED BAYESIAN MODELING

Hierarchical Model for Rat Tumors

Review

$$
\begin{aligned}
n_j &= \text{total number of rats in control group of experiment } j \\
y_j &= \text{number in control group of experiment } j \text{ that develop a tumor} \\
\theta_j &= \text{control-group tumor probability in experiment } j \\
j &= 1, \ldots, 71
\end{aligned}
$$

$$
y_j \mid \theta_j \sim \mathrm{Bin}(n_j, \theta_j)
$$

Seek to model θ_js as if independent from same distribution.

Natural parametric choice for a distribution of probabilities:

$$\mathrm{Beta}(\alpha, \beta) \qquad\qquad \alpha > 0, \quad \beta > 0$$

Continuous, and gives probability 1 to interval $(0, 1)$

Recall density:

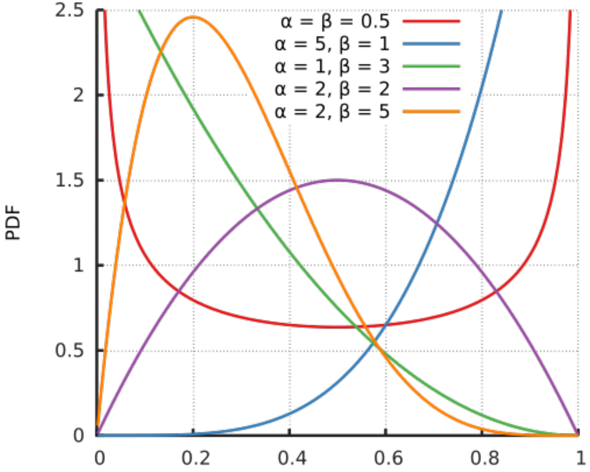
$$p(\theta) \;=\; \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\,\theta^{\alpha-1}\,(1-\theta)^{\beta-1} \qquad 0 < \theta < 1$$

($\mathrm{U}(0, 1)$ is a special case: $\alpha = \beta = 1$)

So try

$$\theta_j \;\sim\; \mathrm{Beta}(\alpha, \beta)$$

Some $\mathrm{Beta}(\alpha, \beta)$ densities:



From: Beta distribution. (2017, May 29). In *Wikipedia, The Free Encyclopedia*. Retrieved June 2, 2017 from `https://en.wikipedia.org/w/index.php?title=Beta_distribution&oldid=782786187`

Model now has two levels:

- Lower level:
$$y_j \mid \theta_j \sim \mathrm{Bin}(n_j, \theta_j)$$
- Higher level:
$$\theta_j \sim \mathrm{Beta}(\alpha, \beta)$$

Called a **hierarchical** (or **multilevel**) model

Call α and β **hyperparameters**.

How to choose them?

- Guess (subjective)
- Use prior information (if available)
- Estimate from data ("empirical Bayes")
- Give them a prior distribution ("hierarchical Bayes")

Following BDA3, we choose the last option.

A prior on hyperparameters is a **hyperprior**.

How to choose a hyperprior for $\alpha > 0$ and $\beta > 0$?

No natural (conjugate) choice

Want something that is

- Convenient to specify and use
- Not too informative

For example:

$$\alpha,\, \beta \;\; \sim \;\; \text{iid Expon}(\lambda) \qquad\qquad \text{some } \lambda > 0$$

so that

$$p(\alpha, \beta) \;\; = \;\; p(\alpha)\, p(\beta) \;\; = \;\; \lambda e^{-\lambda\alpha} \cdot \lambda e^{-\lambda\beta} \qquad \alpha > 0,\, \beta > 0$$

This can be made less informative (flatter) by choosing λ closer to zero.

Alternative suggestion from BDA3 (Sec. 5.3):

$$p\left(\frac{\alpha}{\alpha+\beta},\, (\alpha+\beta)^{-1/2}\right) \;\propto\; 1 \qquad\qquad 0 < \frac{\alpha}{\alpha+\beta} < 1, \quad (\alpha+\beta)^{-1/2} > 0$$

Motivation:

$$\frac{\alpha}{\alpha+\beta} \;=\; \mathrm{E}(\theta_j \mid \alpha, \beta) \;=\; \mu$$

$$(\alpha+\beta)^{-1/2} \;\approx\; \sqrt{\frac{\mathrm{var}(\theta_j \mid \alpha, \beta)}{\mu(1-\mu)}}$$

Note: Improper, but can be shown to give proper posterior

Warning: Using improper hyperpriors can be dangerous (example in BDA3).

Must be able to verify posterior is proper

Independence Assumptions

Natural to assume different experiments are independent (different times, different places, different researchers, ...)

So pairs

$$(y_1, \theta_1), \quad (y_2, \theta_2), \quad \ldots, \quad (y_{71}, \theta_{71})$$

are conditionally independent of each other, given the population of experiments (i.e., given α and β).

Thus θ_js are conditionally independent, given α and β.

We also assume y_j is conditionally independent of α and β, given θ_j.

Can express these relationships through the joint density:

$$
\begin{aligned}
p\big(\{y_j\}, \{\theta_j\}, \alpha, \beta\big) &= p(\alpha, \beta)\, p\big(\{y_j\}, \{\theta_j\} \mid \alpha, \beta\big) \\
&= p(\alpha, \beta) \prod_{j=1}^{71} p(y_j, \theta_j \mid \alpha, \beta) \\
&= p(\alpha, \beta) \prod_{j=1}^{71} p(\theta_j \mid \alpha, \beta)\, p(y_j \mid \theta_j, \alpha, \beta) \\
&= p(\alpha, \beta) \prod_{j=1}^{71} p(\theta_j \mid \alpha, \beta)\, p(y_j \mid \theta_j)
\end{aligned}
$$

Technically more precise to write

$$\{y_j\} \mid \{\theta_j\}, \alpha, \beta \quad \sim \quad \text{indep. Bin}\big(\{n_j\}, \{\theta_j\}\big)$$

For simplicity just write

$$y_j \mid \theta_j \;\; \sim \;\; \text{Bin}(n_j, \theta_j)$$

Conditional independence assumptions are implicit.

Similarly, more precise to write

$$\{\theta_j\} \mid \alpha, \beta \quad \sim \quad \text{iid Beta}(\alpha, \beta)$$

For simplicity just write

$$\theta_j \mid \alpha, \beta \;\; \sim \;\; \text{Beta}(\alpha, \beta)$$

Conditional independence assumption is implicit.

Putting it all together:

$$
\begin{aligned}
y_j \mid \theta_j &\sim \mathrm{Bin}(n_j, \theta_j) \\
\theta_j \mid \alpha, \beta &\sim \mathrm{Beta}(\alpha, \beta) \\
\alpha, \beta &\sim \text{some joint distribution}
\end{aligned}
$$