ADVANCED BAYESIAN MODELING

HIERARCHICAL MODELING FUNDAMENTALS: **EXCHANGEABILITY**

Scenario

Consider data

$$y = (y_1, \dots, y_J)$$

Each y_j has its own parameter θ_j:

$$\theta = (\theta_1, \dots, \theta_J)$$

In rat tumor example, we assumed (conditional) independence of θ_js.

What justifies this?

Consider making a hierarchical model without any data or information about the data (except overall structure and variable types).

If you randomly permuted the observations, would that change the type of hierarchical model you make?

In many cases, no: You have no information to distinguish the observations.

In particular, permuting the θ_js should leave their joint density unchanged.

Exchangeability

Consider (joint) density

$$p(\theta) \;=\; p(\theta_1, \ldots, \theta_J)$$

Random variables $\theta_1, \ldots, \theta_J$ are **exchangeable** if, for any permutation $\pi_1, \ldots, \pi_J$ of indices $1, \ldots, J$,

$$p(\theta_{\pi_1}, \ldots, \theta_{\pi_J}) \;=\; p(\theta_1, \ldots, \theta_J)$$

Note: Implies all have same marginal distribution

What should be exchangeable in a Bayesian model?

Model some random variables as exchangeable (ignoring data values) when they represent the same kind of quantity, and you either

- Have no prior information to make distinctions, or
- Have only unreliable or controversial prior information – choose to ignore, let data decide

Example

Consider rat tumor hierarchical model:

- Model tumor probabilities θ_j as exchangeable: No information to distinguish the experiments (other than the data values), and no reason to believe θ_j depends on group size n_j

- Cannot model data y_j as exchangeable, because n_js differ.

  However, can argue for exchangeability of pairs

  $$(y_1, n_1), \quad \ldots, \quad (y_{71}, n_{71})$$

  (if willing to consider a distribution on n_js)

Relation to Independence

Clearly, iid random variables are exchangeable.

Random variables that are conditionally iid (on some random ϕ) are also exchangeable:

$$\begin{aligned} p(\theta) &= \int p(\theta, \phi)\, d\phi \;=\; \int p(\theta \mid \phi)\, p(\phi)\, d\phi \\ &= \int \left(\prod_{j=1}^{J} p(\theta_j \mid \phi) \right) p(\phi)\, d\phi \end{aligned}$$

For example, θ is a parameter vector and ϕ is a **hyperparameter** with **hyperprior** $p(\phi)$.

Random variables can be exchangeable even when *not* (conditionally) iid.

Nonetheless, usually choose to model exchangeable random variables as conditionally iid, if possible.

(More: BDA3, Sec. 5.2)