Constructing hierarchical models

- Identifying the J experiments/groups that the data come from
- How many repeted observations n_j from each group j, $j=1,\dots,J$
- Find an appropriate sampling distribution for y_{ij}, the i'th observation from the j'th group, parametrised by a group-specific parameter (vector) θ_j

$$y_{ij} \sim p(y_{ij} \mid \theta_j) \quad , \quad i=1,\dots,n_j, \quad j=1,\dots,J$$

For example $y_{ij} \sim N(\mu_j, \sigma^2)$

If we have covariate-information x_{ij} for the individual observations

$$y_{ij} \sim p(y_{ij} \mid \theta_j, x_{ij})$$

i.e. $(y_{ij}, \theta_j, x_{ij})$ are exchangeable

Example Rats: $y_j \sim Bin(\theta_j, n_j)$, i.e. (y_j, θ_j, n_j) are exchangeable

- If no other information is available to distinguish the experiments/groups, then it is appropriate to assume that $p(\theta_1,\dots,\theta_J)$ is invariant to permutation of the indexes $1,\dots,J$

The simplest form: Assume the θ_j's to be iid from a prior distribution parametrised by a hyperparameter (vector) ϕ, i.e.

$$p(\theta_1,\dots,\theta_J \mid \phi) = \prod_{j=1}^{J} p(\theta_j \mid \phi)$$

okt 27-12:13

- More advanced, if we have covariate information z_j on each group $j=1,\dots,J$, then it is appropriate to model the θ_j's as conditionally independent given covariate information, i.e.

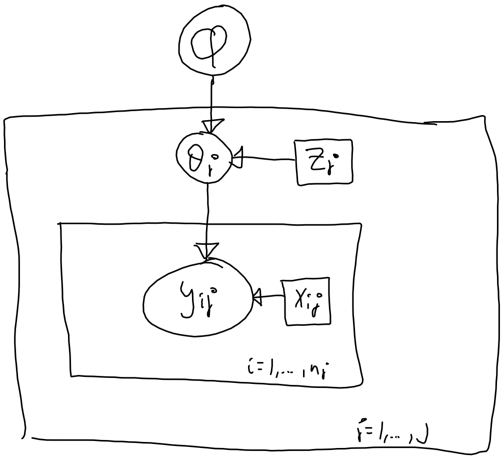
$$p(\theta_1,\dots,\theta_J \mid \phi, z_1,\dots,z_J) = \prod_{j=1}^{J} p(\theta_j \mid \phi, z_j)$$

Hence, the pairs (θ_j, z_j) are exchangeable

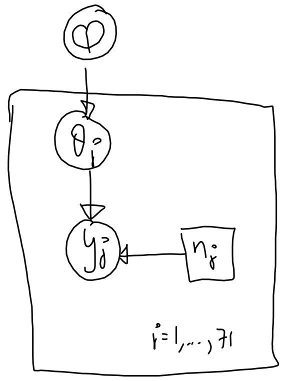
As long as all information available to distinguish the groups is incorporated into z_j, $j=1,\dots,J$, then such a model is appropriate.

- Could also be appropriate to add another layer, e.g. in the Rats example we knew that specific groups of experiments where done in 5 different labs.
- Hyperprior distribution for ϕ
    - If little prior knowledge available, use a diffuse prior
        - BUT be careful with improper prior distributions, can result in improper posterior distributions
            - ↳ The more layers and parameters, the more difficult it gets to that the posterior distribution is proper (can be hidden)

okt 27-12:30



Rats:



okt 27-12:43

Posterior predictive inference

Two different kinds of posterior predictive quantities:

(1) Predict $\tilde{y}_j$ for some j in $\{1,\dots,J\}$

- Draw $\tilde{y}_j$ based on posterior draws θ_j

(2) Predict $\tilde{y}_{\tilde{j}}$ for a new group $\tilde{j} \notin \{1,\dots,J\}$

- Draw $\tilde{\theta}_{\tilde{j}}$ based on posterior draws of ϕ
- Draw $\tilde{y}_{\tilde{j}} \mid \tilde{\theta}_{\tilde{j}}$

okt 27-12:49