# Hierarchical models

- In the non-hierarchical models we simply had
  - Conditional density of data $\pi$(y | parameters).
  - Prior density of parameters $\pi$(parameters).

  - to get posterior $\pi$(parameters | y).

# Hierarchical models

- In the non-hierarchical models
  - As a default, the prior is often ’noninformative’.

    → Can expect posterior mode to resemble maximum likelihood estimate, since we basically ”normalize the likelihood function” to become a probability density:

    $\pi$(parameters|y)=$\pi$(y| parameters)×”1”/const.

    → not utilizing the full potential of Bayesian models, unless we elicit informative priors?

# Hierarchical models

- In the non-hierarchical models
  - Informative priors can be laborious to get, although may be essential ingredient in some problems. (e.g. expert opinions).
  - ”Empirical Bayes”: draw the informative priors from data
    - ”Prior” data should be independent of the ”actual” data, to avoid using same data twice.

# Hierarchical models

- With hierarchical data (groups within groups)
  - Natural idea to borrow strength from rest of the groups, to support weak data in some groups → add hyperprior to the model to make it hierarchical.
  - Accomplishes the effect of informative prior, but over whole hierarchical structure, based on **conditional independency**.

  → extremely versatile approach to many "hard problems" where data are unbalanced, partially missing, etc, but where the total information can help "filling the gaps".

  → Bayesian "information synthesis".