# STAT 530: Hierarchical Models

Mar. 8, 2010

## Generic Hierarchical Model (indices backward?)

$$
\begin{aligned}
y_{i,j} &\sim p(y|\phi_j) \text{ (independently across } i,j) \\
\phi_1, \dots, \phi_m &\overset{\text{iid}}{\sim} p(\phi|\psi) \\
\psi &\sim p(\psi)
\end{aligned}
$$

Exchangeability ideas again.

Posterior: $(\phi_1, \dots, \phi_m, \psi|y)$

Contexts for $(i,j)$?

## Normal Case

$$
\begin{aligned}
y_{i,j} &\sim N(\theta_j, \sigma^2) \\
\theta_1, \dots, \theta_m &\overset{\text{iid}}{\sim} N(\mu, \tau^2) \\
\mu, \tau^2, \sigma^2 &\sim p(\mu)p(\tau^2)p(\sigma^2)
\end{aligned}
$$

with $\mu \sim N(\mu_0, \gamma^2)$.

For now, focus on special case of $\sigma^2, \tau^2$ known.

Implications of assuming $\tau^2 = 0$?

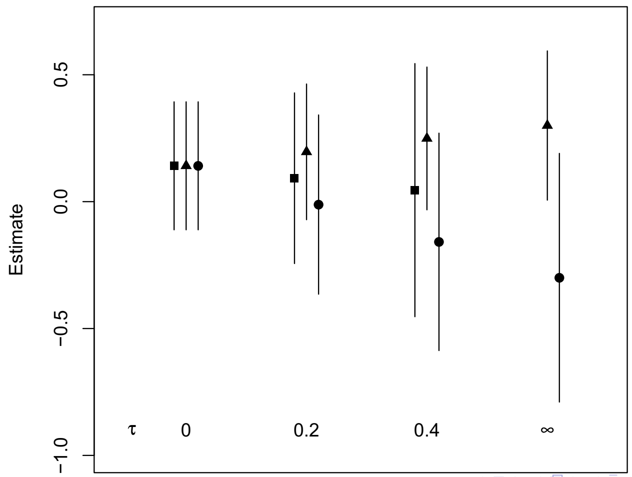
Implications of assuming $\tau^2 = \infty$?

## Structure of Posterior

$(\theta_j|\mu, \sigma^2, \tau^2, y)$

$(\mu|\sigma^2, \tau^2, y)$

## Ex.: $\bar{y}_1 = 0.3$, $\bar{y}_2 = -0.3$, $\sigma = 0.75$, $n_1 = 25$, $n_2 = 9$ (and $\gamma^2$ large)




## $\tau^2$ controls 'shrinkage' - let data decide?

*IG* priors for $\sigma^2$, $\tau^2$ yield IG posterior full conditionals.

So easy to Gibbs sample (can update $(\theta, \mu)$ together, or as separate blocks).

## Comments on Text Example

$m = 100$ schools

$y_{i,j}$ is math score for $i$-th student in $j$-th school.

$n_j$ ranges from 5 to 32.

Vague prior for $\tau^2$: $\eta_0 = 1, \tau_0^2 = 100$.
Posterior very concentrated in relation to this.
Data are speaking to how much shrinkage should occur.

Plotting $\hat{\theta}_j = E(\theta_j|y)$ against $\bar{y}_j$ illustrates shrinkage effect.

## Comments, continued

Tendency for $|\bar{y}_j - \hat{\theta}_j|$ to be smaller when $n_j$ is larger - makes sense.

Shrinkage can reverse order.

E.g., possible that $\bar{y}_j > \bar{y}_k$ but $\hat{\theta}_j < \hat{\theta}_k$.

Again makes sense, but tough sell to non-statisticians?