StatMod2 - Hierarchical Models - Exercises 4

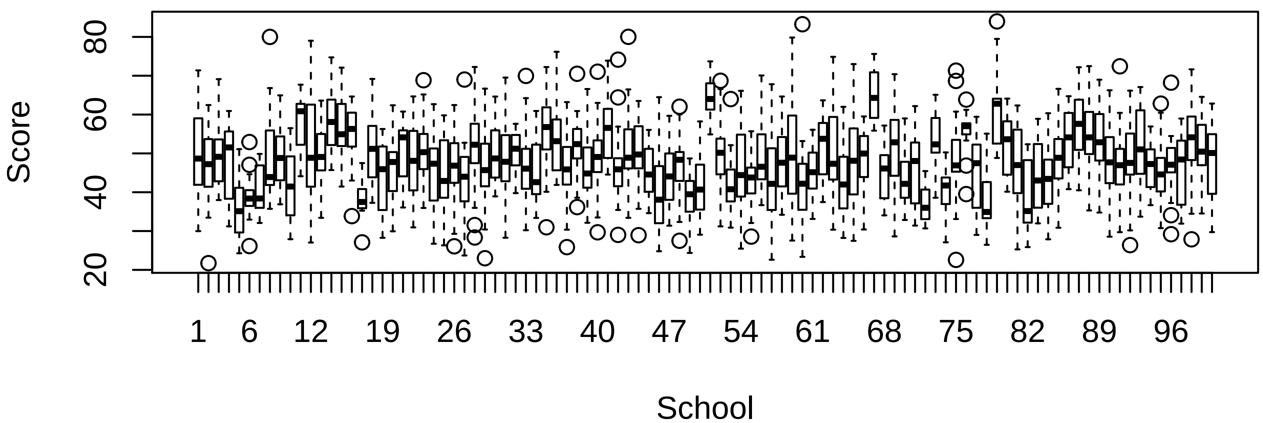
Maurice Diesendruck

April 14, 2015

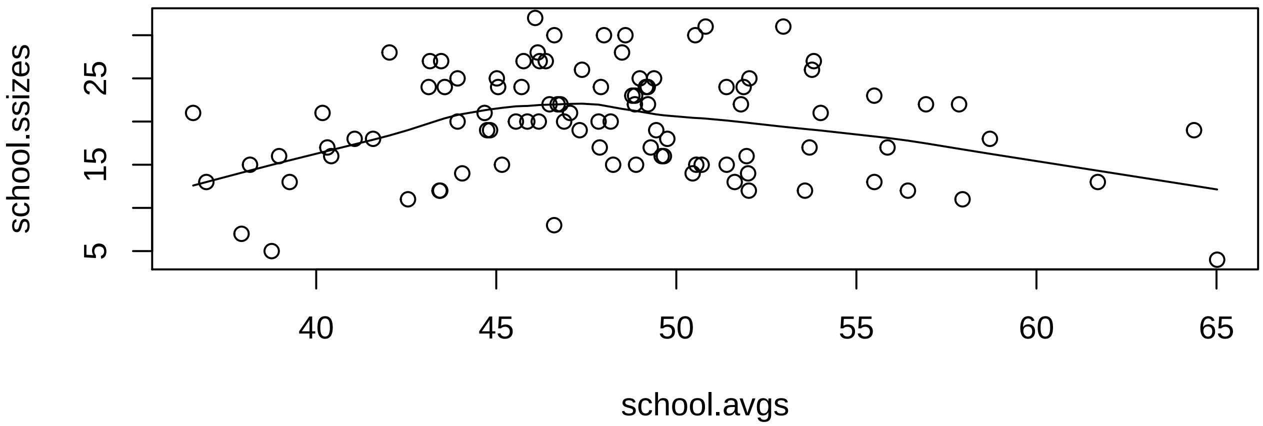
1 School Averages and Sample Size

Larger samples tend to ”smooth” out extreme scores, so small samples are more likely to be extreme.

```
data <- read.csv("mathtest.csv")
attach(data)
boxplot(mathscore ~ school, xlab="School", ylab="Score")
```




```
school.avgs <- aggregate(data, list(school=school), mean)[,3]
school.ssizes <- aggregate(data, list(school=school), length)[,3]
plot(cbind(school.avgs, school.ssizes))
fit <- lowess(cbind(school.avgs, school.ssizes))
lines(fit)
```



2 Normal Hierarchical Model with Gibbs Sampling

2.1 Model

For school $i = 1, \cdots, p$; student $j = 1, \cdots, n_i$; and $\sum n_i = n$; let $a = b = c = d = 1$, and note that for this data, $p = 100$ and $n = 1993$:

$$
\begin{aligned}
y_{ij} &\sim N(\theta_i, \sigma^2) \\
\theta_i &\sim N(\mu, \tau^2) \\
\mu &\sim Unif(40, 60) \\
\sigma^2 &\sim InvGa(a, b) \\
\tau^2 &\sim InvGa(c, d)
\end{aligned}
$$

Or alternatively,

$$
\begin{aligned}
\bar{y}_i &\sim N(\theta_i, \frac{\sigma^2}{n_i}) \\
\theta_i &\sim N(\mu, \tau^2) \\
\mu &\sim Unif(40, 60) \\
\sigma^2 &\sim InvGa(a, b) \\
\tau^2 &\sim InvGa(c, d)
\end{aligned}
$$

$$
N(\bar{\boldsymbol{y}}|\boldsymbol{\theta}, \sigma^2 diag(1/n_i))N(\boldsymbol{\theta}|\boldsymbol{\mu}, \tau^2 I_p)Unif(\mu|40, 60)Ga(\frac{1}{\sigma^2}|a, b)Ga(\frac{1}{\tau^2}|c, d)
$$

2.2 Joint and Posterior Distributions

Let $\boldsymbol{\nu} = (\boldsymbol{\theta}, \boldsymbol{\mu}, \sigma^2, \tau^2)$; $\boldsymbol{\theta} = \theta_1, \cdots, \theta_p$; $\boldsymbol{\sigma^2_*} = \sigma^2 diag(1/n_i)$; $\bar{\boldsymbol{y}}$ be the px1 vector of mean math scores for all schools; and note that $\boldsymbol{\mu}$ is a px1 vector with a single value, μ.

$$
\begin{aligned}
P(\boldsymbol{\nu}|\boldsymbol{y}) \propto\ & P(\boldsymbol{y}|\boldsymbol{\nu})P(\boldsymbol{\nu}) \\
&= N(\bar{\boldsymbol{y}}|\boldsymbol{\theta}, \boldsymbol{\sigma^2_*})N(\boldsymbol{\theta}|\boldsymbol{\mu}, \tau^2 I_p)Unif(\mu|40, 60)Ga(\frac{1}{\sigma^2}|a, b)Ga(\frac{1}{\tau^2}|c, d) \\
&= \Big(\frac{1}{\boldsymbol{\sigma^2_*}}\Big)^{p/2} exp\Big(-\frac{1}{2}\frac{(\bar{\boldsymbol{y}} - \boldsymbol{\theta})'(\bar{\boldsymbol{y}} - \boldsymbol{\theta})}{\boldsymbol{\sigma^2_*}}\Big) \\
&\quad \text{x } \Big(\frac{1}{\tau^2}\Big)^{p/2} exp\Big(-\frac{1}{2}\Big(\frac{(\boldsymbol{\theta} - \boldsymbol{\mu})'(\boldsymbol{\theta} - \boldsymbol{\mu})}{\tau^2}\Big)\Big) \\
&\quad \text{x } \Big(\frac{1}{\sigma^2}\Big)^{a-1} exp\Big(-b\Big(\frac{1}{\sigma^2}\Big)\Big) \\
&\quad \text{x } \Big(\frac{1}{\tau^2}\Big)^{c-1} exp\Big(-d\Big(\frac{1}{\tau^2}\Big)\Big)
\end{aligned}
$$

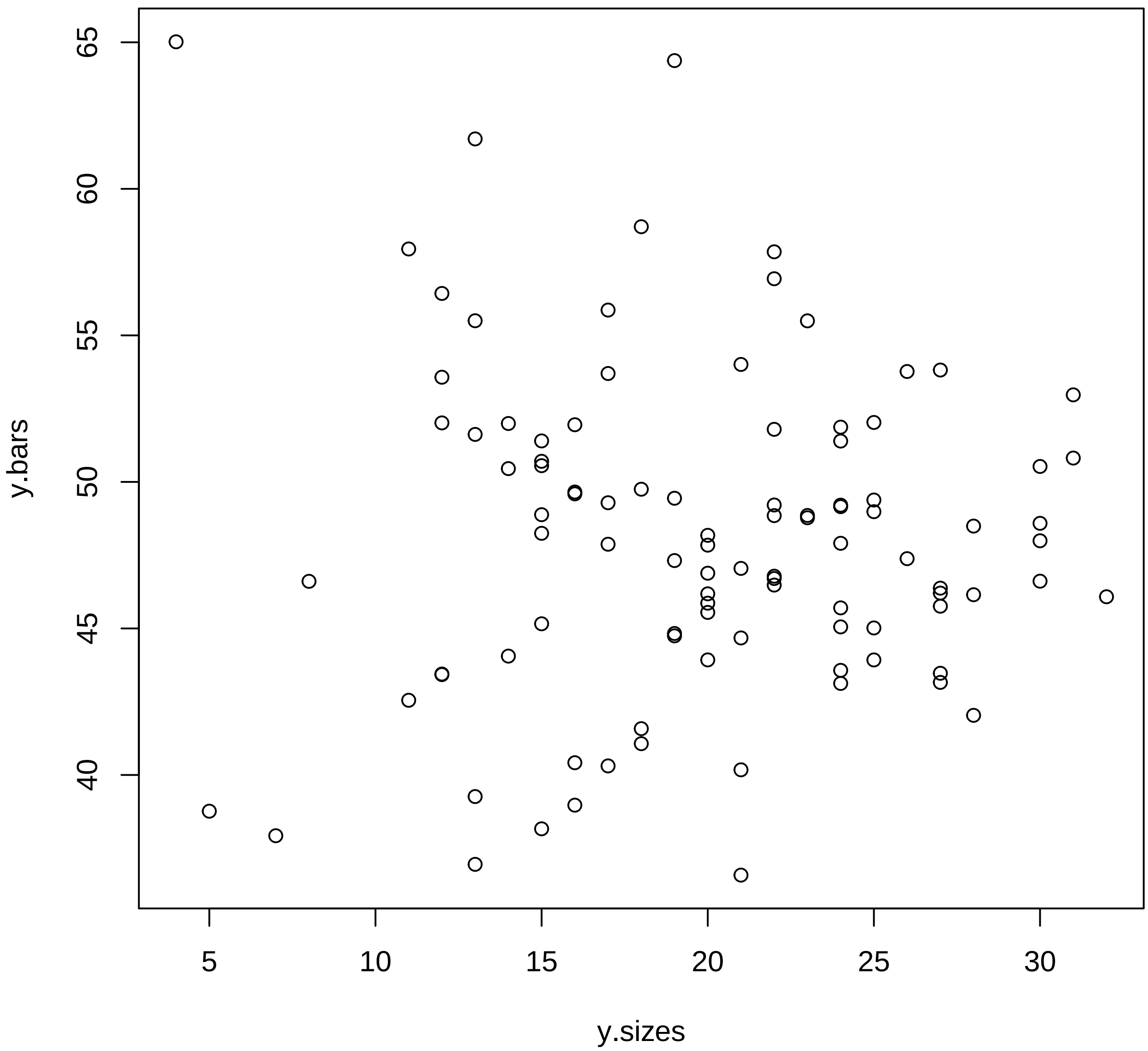
Conditional distributions are as follows:

$$
\begin{aligned}
P(\boldsymbol{\theta}|\bar{\boldsymbol{y}},\boldsymbol{\mu},\sigma^2,\tau^2) \propto\ & exp\Big(-\frac{1}{2}\Big(\frac{(\bar{\boldsymbol{y}}-\boldsymbol{\theta})'(\bar{\boldsymbol{y}}-\boldsymbol{\theta})}{\boldsymbol{\sigma_*^2}}+\frac{(\boldsymbol{\theta}-\boldsymbol{\mu})'(\boldsymbol{\theta}-\boldsymbol{\mu})}{\tau^2}\Big)\Big)\\
&= exp\Big(-\frac{1}{2}\Big(\frac{\bar{\boldsymbol{y}}'\bar{\boldsymbol{y}}}{\boldsymbol{\sigma_*^2}}-\frac{\bar{\boldsymbol{y}}'\boldsymbol{\theta}}{\boldsymbol{\sigma_*^2}}-\frac{\boldsymbol{\theta}'\bar{\boldsymbol{y}}}{\boldsymbol{\sigma_*^2}}+\frac{\boldsymbol{\theta}'\boldsymbol{\theta}}{\boldsymbol{\sigma_*^2}}\Big)+\Big(\frac{\boldsymbol{\theta}'\boldsymbol{\theta}}{\tau^2}-\frac{\boldsymbol{\theta}'\boldsymbol{\mu}}{\tau^2}-\frac{\boldsymbol{\mu}'\boldsymbol{\theta}}{\tau^2}+\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\sigma^2}\Big)\Big)\\
&\sim\ N\Bigg(\Big(\frac{I_n}{\boldsymbol{\sigma_*^2}}+\frac{I_n}{\tau^2}\Big)^{-1}\Big(\frac{\bar{\boldsymbol{y}}}{\boldsymbol{\sigma_*^2}}-\frac{\boldsymbol{\mu}}{\tau^2}\Big),\Big(\frac{I_n}{\boldsymbol{\sigma_*^2}}+\frac{I_n}{\tau^2}\Big)^{-1}\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
P(\sigma^2|\boldsymbol{y},\boldsymbol{\theta},\tau^2) \propto\ & \Big(\frac{1}{\sigma^2}\Big)^{n/2} exp\Big(-\frac{1}{2}\Big(\frac{(\bar{\boldsymbol{y}}-\boldsymbol{\theta})'(\bar{\boldsymbol{y}}-\boldsymbol{\theta})}{\boldsymbol{\sigma_*^2}}\Big)\Big)\\
&\quad \text{x}\ \Big(\frac{1}{\sigma^2}\Big)^{a-1} exp\Big(-b\Big(\frac{1}{\sigma^2}\Big)\Big)\\
&= \Big(\frac{1}{\sigma^2}\Big)^{n/2+a-1} exp\Big(-\Big(\frac{(\bar{\boldsymbol{y}}-\boldsymbol{\theta})'(\bar{\boldsymbol{y}}-\boldsymbol{\theta})}{2diag(1/n_i)}+b\Big)\Big(\frac{1}{\sigma^2}\Big)\Big)\\
&\sim\ Ga\Big(n/2+a,\Big(\frac{(\bar{\boldsymbol{y}}-\boldsymbol{\theta})'(\bar{\boldsymbol{y}}-\boldsymbol{\theta})}{2diag(1/n_i)}+b\Big)\Big)
\end{aligned}
$$

$$
\begin{aligned}
P(\tau^2|\boldsymbol{y},\boldsymbol{\theta},\sigma^2) \propto\ & \Big(\frac{1}{\tau^2}\Big)^{p/2} exp\Big(-\frac{1}{2}\Big(\frac{(\boldsymbol{\theta}-\boldsymbol{\mu})'(\boldsymbol{\theta}-\boldsymbol{\mu})}{\tau^2}\Big)\Big)\\
&\quad \text{x}\ \Big(\frac{1}{\tau^2}\Big)^{c-1} exp\Big(-d\Big(\frac{1}{\tau^2}\Big)\Big)\\
&= \Big(\frac{1}{\tau^2}\Big)^{p/2+c-1} exp\Big(-\Big(\frac{(\boldsymbol{\theta}-\boldsymbol{\mu})'(\boldsymbol{\theta}-\boldsymbol{\mu})}{2}+d\Big)\Big(\frac{1}{\tau^2}\Big)\Big)\\
&\sim\ Ga\Big(p/2+c,\Big(\frac{(\boldsymbol{\theta}-\boldsymbol{\mu})'(\boldsymbol{\theta}-\boldsymbol{\mu})}{2}+d\Big)\Big)\\
P(\boldsymbol{\mu}|\boldsymbol{y},\boldsymbol{\theta},\sigma^2,\tau^2) \propto\ & exp\Big(-\frac{1}{2}\Big(\frac{(\boldsymbol{\theta}-\boldsymbol{\mu})'(\boldsymbol{\theta}-\boldsymbol{\mu})}{\tau^2}\Big)\Big)\\
&\sim\ N(\bar{\boldsymbol{\theta}},\tau^2)
\end{aligned}
$$

2.3 Gibbs Sampler




```
## Error in eval(expr, envir, enclos):  object 'ybars' not found
```

2.4 Shrinkage

In general, the smaller the sample size, the more extreme the sample mean, and the larger the shrinkage factor.