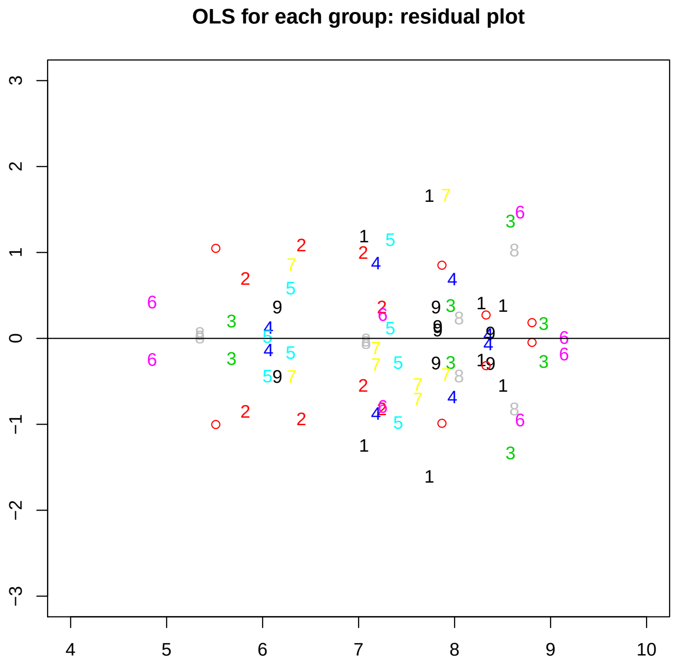
Introduction to Bayesian Data Analysis

Tutorial 11 - Solutions

(1) Problem 11.2 (Hoff)

(a) ‘




```
> theta<-apply(BETA.LS,2,mean)
> Sigma<-cov(BETA.LS)
> s2<-mean(S2.LS)
```

Our ad-hoc estimates are as follows

$$\hat{\boldsymbol{\theta}} = (2.86875, 1.85485, -0.15925);\ \hat{\Sigma} = \begin{pmatrix} 2.0012 & -0.6932 & 0.04431 \\ -0.6932 & 0.2756 & -0.02074 \\ 0.0443 & -0.0207 & 0.00197 \end{pmatrix}$$

$\hat{\sigma^2} = 0.788$

(b) See R code for Gibbs sampler algorithm

(c) `BETA.PM<-BETA.ps/(S/thin) #posterior expectation for beta_j`

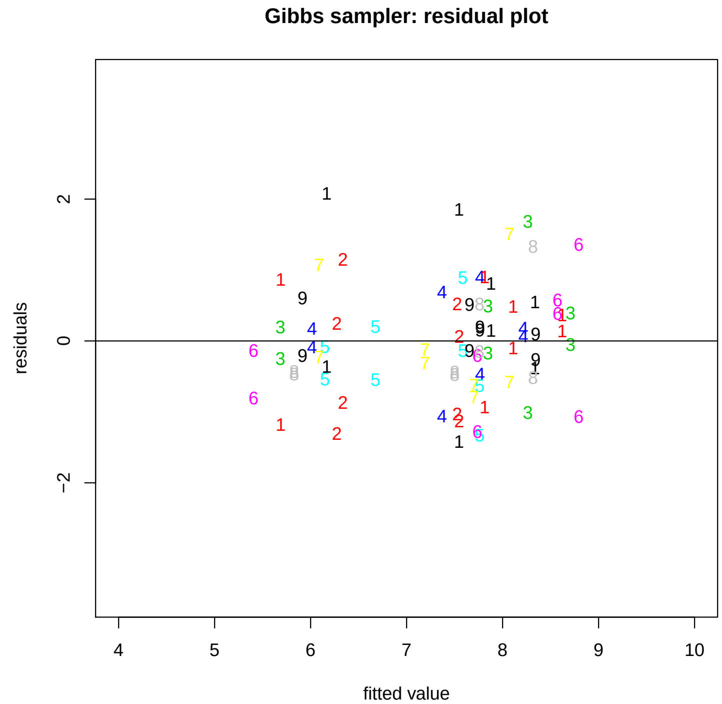
Group	intercept	x	x2
1	3.21	1.80	-0.157
2	3.89	1.53	-0.154
3	2.22	2.06	-0.163
4	2.97	1.85	-0.162
5	3.46	1.66	-0.157
6	1.80	2.12	-0.160
7	3.23	1.74	-0.156
8	2.53	1.99	-0.172
9	2.80	1.88	-0.159
10	2.25	2.05	-0.165

The plot of the posterior expectations of the residuals versus fitted values is less spread out along the x-axis. This is because of the shrinkage effect from fitting a hierarchical model.

(d) Marginal posterior density plots of the elements of Σ are peaked around a value of zero, indicating that that there does not seem to be much variation in slopes or intercepts across groups.

(e)
```
> fitted<-NULL
> for(j in 1:m) {
+  fitted<-rbind(fitted,c(X[plot==j,]%*%BETA.PM[j,])) }
> c(mean(fitted[,1:2]),mean(fitted[,3:4]),
       mean(fitted[,5:6]),mean(fitted[,7:8]))
[1] 5.93 7.74 8.27 7.52
> xmax<-c(1,6,36) #max at x=6
> xmax.pred<-xmax%*%t(BETA.pp)
> quantile(xmax.pred,probs=c(0.026,0.975))
 2.6% 97.5%
 7.08  9.22
```

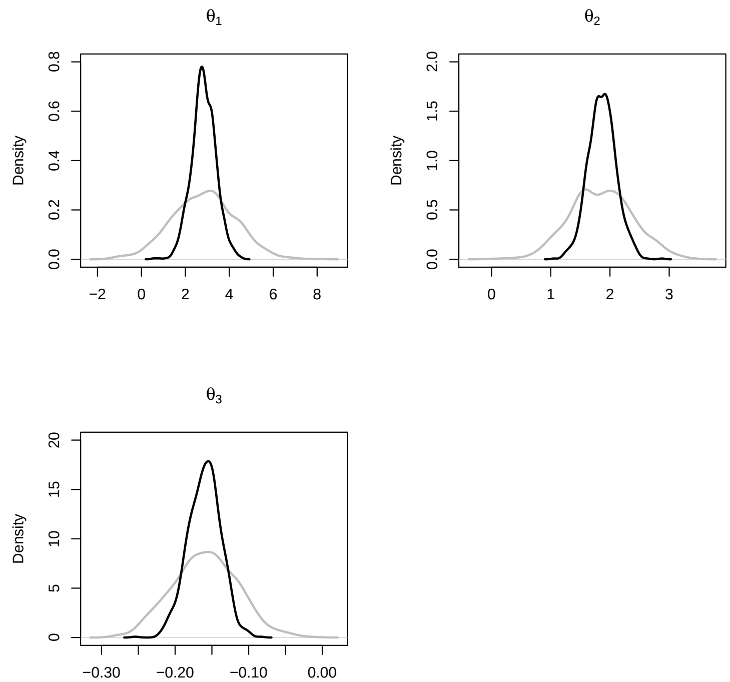
The value of x that maximises expected yield is 6, and a 95% predictive interval for the yield of a randomly sampled plot with $x = 6$ is (7.08,9.22)



(2) (a)

$$
\begin{aligned}
p(\sigma_0^2|\boldsymbol{\beta}_1,....,\boldsymbol{\beta}_m,\sigma_1^2,...,\sigma_m^2,X,\nu_0,\mathbf{y},\boldsymbol{\theta},\Sigma) &\propto p(\sigma_0^2)\prod_{j=1}^{m} p(\sigma_j^2|\nu_0,\sigma_0^2) \\
&= \sigma_0^2\exp(-2\sigma_0^2)\times(\sigma_0^2)^{\nu_0 m/2}\exp\left(-\frac{\nu_0\sigma_0^2}{2}\sum_{j=1}^{m}1/\sigma_j^2\right) \\
&= (\sigma_0^2)^{1+\nu_0 m/2}\exp\left(-\sigma_0^2(2+\nu_0/2\sum_{j=1}^{m}1/\sigma_j^2)\right) \\
&= (\sigma_0^2)^{(2+\nu_0 m/2)-1}\exp\left(-\sigma_0^2(2+\nu_0/2\sum_{j=1}^{m}1/\sigma_j^2)\right)
\end{aligned}
$$

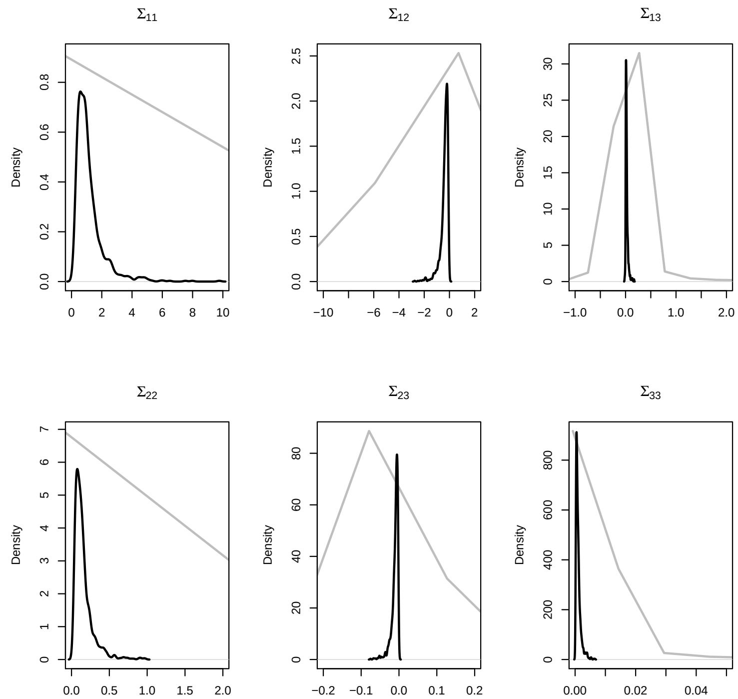
Therefore we see that the full conditional distribution of σ_0^2 is Gamma$(2+\nu_0 m/2\ ,\ 2+\nu_0/2\sum_{j=1}^{m}1/\sigma_j^2)$.



(b)

$$
\begin{aligned}
& p(\sigma_j^2|\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_m, \sigma_0^2, X, \nu_0, \mathbf{y}, \boldsymbol{\theta}, \Sigma) \\
& \propto p(\sigma_j^2|\sigma_0^2, X, \nu_0) \prod_{i=1}^{n_j} p(y_{i,j}|X_j, \boldsymbol{\beta}_j, \sigma_j^2) \\
& = (\sigma_j^2)^{-(\nu_0/2+1)} \exp\left(-\frac{\nu_0\sigma_0^2}{2\sigma_j^2}\right) (\sigma_j^2)^{-n_j/2} \exp\left(-\frac{1}{2\sigma_j^2}\sum_{i=1}^{n_j}(y_{ij} - X_j^T\boldsymbol{\beta_j})^2\right) \\
& = (\sigma_j^2)^{-(\nu_0/2+n_j/2+1)} \exp\left(-\frac{\nu_0\sigma_0^2 + \sum_{i=1}^{n_j}(y_{ij} - X_j^T\boldsymbol{\beta_j})^2}{2\sigma_j^2}\right)
\end{aligned}
$$

Therefore we see that the full conditional distribution of σ_j^2 is Inverse-Gamma $\left(\nu_0/2 + n_j/2, \frac{\nu_0\sigma_0^2+\sum_{i=1}^{n_j}(y_{ij}-X_j^T\boldsymbol{\beta_j})^2}{2}\right)$.



(c)

$$
\begin{aligned}
&p(\boldsymbol{\beta}_j|\sigma_1^2,...,\sigma_m^2,X,\sigma_0^2,\nu_0,\mathbf{y},\boldsymbol{\theta},\Sigma)\\
&\propto p(\boldsymbol{\beta}_j|\boldsymbol{\theta},\Sigma)\prod_{i=1}^{n_j}p(y_{i,j}|X_j,\boldsymbol{\beta}_j,\sigma_j^2)\\
&\propto \exp\left(-\frac{1}{2}(-2\boldsymbol{\beta}_j^T\Sigma^{-1}\boldsymbol{\theta}+\boldsymbol{\beta}_j^T\Sigma^{-1}\boldsymbol{\beta}_j)\right)\times\exp\left(-\frac{1}{2}(-2\boldsymbol{\beta}_j^TX_j^T\mathbf{y}_j/\sigma_j^2+\boldsymbol{\beta}_j^TX^TX\boldsymbol{\beta}_j/\sigma_j^2)\right)\\
&=\exp\left\{\boldsymbol{\beta}_j^T(\Sigma^{-1}\boldsymbol{\theta}+X^T\mathbf{y}/\sigma_j^2)-\frac{1}{2}\boldsymbol{\beta}_j^T(\Sigma^{-1}+X^TX/\sigma_j^2)\Sigma^{-1}(\boldsymbol{\theta}+X^T\mathbf{y}/\sigma_j^2)\right\}
\end{aligned}
$$

Therefore we see that the full conditional distribution of $\boldsymbol{\beta}_j^2$ is $\text{MVN}\big((\Sigma^{-1}+X^TX/\sigma_j^2)^{-1}(\Sigma^{-1}\boldsymbol{\theta}+X^T\mathbf{y}/\sigma_j^2),(\Sigma^{-1}+X^TX/\sigma_j^2)^{-1}\big)$.

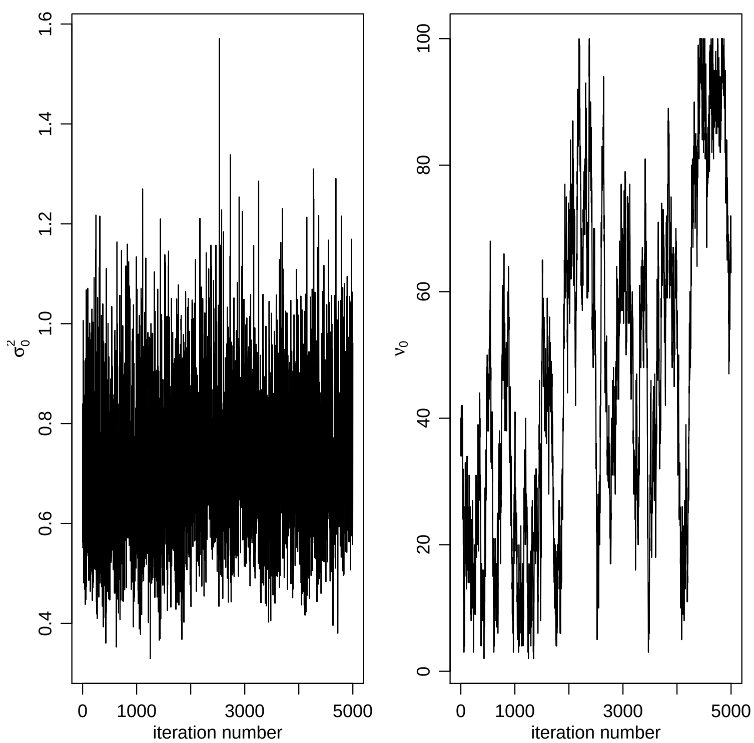
(d)

$$\frac{p(\nu_0^*|\sigma_0^2,\sigma_1^2,...,\sigma_m^2)}{p(\nu_0^{(s)}|\sigma_0^2,\sigma_1^2,...,\sigma_m^2)}$$

$$= \frac{\frac{\left(\nu_0\sigma_0^2/2\right)^{\nu_0^* m/2}}{\Gamma(\nu_0^*/2)}(\prod_{j=1}^m \sigma_j^2)^{-(\nu_0^*/2+1)}\exp\left(-\nu_0^*\sigma_0^2\sum_{j=1}^m 1/\sigma_j^2\right)}{\frac{\left(\nu_0^{(s)}\sigma_0^2/2\right)^{\nu_0^{(s)} m/2}}{\Gamma(\nu_0^{(s)}/2)}(\prod_{j=1}^m \sigma_j^2)^{-(\nu_0^{(s)}/2+1)}\exp\left(-\nu_0^{(s)}\sigma_0^2\sum_{j=1}^m 1/\sigma_j^2\right)}$$

$$= \frac{\frac{\left(\nu_0\sigma_0^2/2\right)^{\nu_0^* m/2}}{\Gamma(\nu_0^*/2)}}{\frac{\left(\nu_0^{(s)}\sigma_0^2/2\right)^{\nu_0^{(s)} m/2}}{\Gamma(\nu_0^{(s)}/2)}}(\prod_{j=1}^m \sigma_j^2)^{(\nu_0^{(s)}/2+1)-(\nu_0^*/2+1)}\exp\left((\nu_0^{(s)}-\nu_0^*)\sigma_0^2\sum_{j=1}^m 1/\sigma_j^2\right)$$

(e) The plot for σ_0^2 shows good mixing. The MCMC chain for ν_0 moves more slowly, and appears cutoff at a value of 100 as this was the maximum possible value allowed in the Markov chain.



(f) There are two modes in the posterior density plot for ν_0, around the values of $\nu_0 = 60$ and $\nu_0 = 20$. $\nu_0 = \infty$ corresponds to the equal variance model, so we do have some evidence that the variances differ across groups.