Homework 4 – Solutions

In several problems we will use the following usefull lemmas. See, for example, DeGroot for proofs (of these or similar results). Also, all three results are easy to show by algebra.

Lemma 1: If $p(y|x) = N(Ax + a, \Sigma)$ and $p(x) = N(\mu, T)$. Assume that A is of *full row rank* (i.e. $w = Ax \sim N(A\mu, ATA'))$. Then

$$p(x) = N(A\mu + a, \Sigma + ATA').$$

Lemma 2: If $x \sim N(m, S)$ is a multivariate normal, then $w = Ax + b$ is $w \sim N(Am + b, ASA')$.

Lemma 3: If $x \sim N(m, S)$ is multivariate normal, and we partition $x = (x_1, x_2)$, $m = (m_1, m_2)$,

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

then the marginal in x_1 is $x_1 \sim N(m_1, S_{11})$. And the conditional $p(x_2|x_1) = N(m_2 - S_{12}S_{11}^{-1}(x_1 - \mu_1), S_{22} - S_{12}S_{11}S_{21})$.

5.2 (a)

$$\begin{aligned} p(\theta_i|\mu_i) &= N(\theta_i|\mu_i, 1), \quad i = 1, \ldots, 2J, \text{ independently} \\ \mu_i &= \begin{cases} -1 & \text{with prob } 1/2, \\ 1 & \text{with prob } 1/2, \end{cases}, \quad \sum \mu_i = 0. \end{aligned}$$

The model is invariant under arbitrary permutations of the indices $(1, \ldots, 2J)$, i.e. $(\theta_1, \ldots, \theta_{2J})$ are finitely exchangeable.

(b) Conditional on $\theta_1, \ldots, \theta_{2J}$ then θ_{2J} we – almost know μ_{2J}, i.e. θ_{2J} cannot be independent of $\theta_{1,\ldots,2J}$. (c) There is no last element if $J \to \infty$

5.7. Define $u = \log(\alpha/\beta)$, $v = \log(\alpha+\beta)$. Then $\beta = e^v/(1+e^u)$ and $\alpha = e^(u+v)/(1+e^u)$ and by the change of variables formulas:

$$p_{\alpha,\beta}(\alpha, \beta) = p_{u,v}[\log(\alpha/\beta), \log(\alpha + \beta)] \cdot \frac{1}{\alpha\beta}.$$

(a) $p_{u,v}(u, v) = 1$, then $p_{\alpha,\beta}(\alpha, \beta) = 1/(\alpha\beta)$ and

$$p(\alpha, \beta|y) = \frac{1}{\alpha\beta} \underbrace{\prod_{j=1}^{J} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)}}_{F1} \underbrace{\prod_{j=1}^{J} \frac{\Gamma(\alpha + n_i - y_i)\Gamma(\beta + y_i)}{\Gamma(\alpha + \beta + n_i)\Gamma(\beta)}}_{F2}.$$

Note that $F1 \geq 1$. For all $y = (y_1, \ldots, y_J)$ with $y_i < n_i$ we find $F2 \to \prod[\Gamma(n_i - y_i)\Gamma(\beta + y_i)]/[\Gamma(\beta + n_i)\Gamma(\beta)]$ as $\alpha \to 0$. Hence there exist c, α_0 s.t. $F2 > c$ for all $\alpha < \alpha_0$. Thus, for $\alpha < \alpha_0$ we have:

$$p(\alpha, \beta|y) \geq \frac{1}{\alpha\beta} c.$$

Therefore $\int_0^{\alpha_0} p(\alpha, \beta|y) d\alpha = \infty$ and $\int p(\alpha, \beta|y) d\alpha d\beta = \infty$.

(b) $r = \alpha/(\alpha + \beta)$, $s = (\alpha + \beta)^{-1/2}$ defines a function $f(\alpha, \beta) = (r, s)$. By the change of variable formula $p_{\alpha,\beta}(\alpha, \beta) = p_{r,s}[f(\alpha, \beta)]|J|$, where

$$J = det \begin{bmatrix} dr/d\alpha & dr/d\beta \\ ds/d\alpha & ds/d\beta \end{bmatrix} = -(\alpha + \beta)^{-5/2},$$

i.e. $p(\alpha, \beta) \propto (\alpha+\beta)^{-5/2}$ and $p_{u,v}(u, v) \propto \alpha\beta(\alpha+\beta)^{-5/2}$.
(c) Denote with $B(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1} = [\Gamma(a)\Gamma(b)]/\Gamma(a+b)$ the beta function. Note that $\lim_{a\to 0} B(a, b) = \lim_{b\to 0} B(a, b) = \infty$ for any b and a, respectively. Also, $B(a, b) > B(c, d)$ for $c \geq a, d \geq b$. Both statments are easily verified by considering the integral defining $B(a, b)$. The posterior $p(\alpha, \beta|y)$ (from (5.8)) can be written as:

$$p(\alpha, \beta|y) \propto (\alpha+\beta)^{-5/2} \prod_{j=2}^{J} \underbrace{B(\alpha+y_j, \beta+n_j-y_j)/B(\alpha, \beta)}_{F_i}.$$

5.9 Let $\eta_i = logit(\theta_i)$, i.e. $\theta_i = exp(\eta_i)/(1+exp(\eta_i))$.

$$\begin{aligned} y_i &\sim Bin(n_i, \theta_i),\ \theta_i = e^{\eta_i}/(1+e^{\eta_i}) \\ \eta_i &\sim N(\mu, \tau^2), \\ p(\mu, \tau) &\propto 1/\tau. \end{aligned}$$

(a)

$$p(\eta, \mu, \tau|y) \propto 1/\tau \prod N(\eta_i; \mu, \tau^2) \cdot \prod \theta_i^{y_i}(1-\theta_i)^{n_i-y_i}.$$

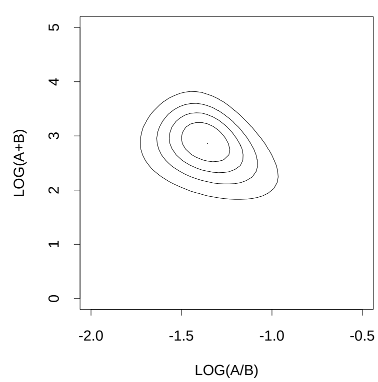
Transform to $p(\theta, \mu, \tau|y)$ by multiplying with the Jacobian $|J| = \prod 1/(\theta_i \cdot (1-\theta_i))$ if you wish.
(b) Integrating out $\eta_1, \ldots, \eta_J$ would amount to computing the expected value of $f(\eta) = \prod \theta_i^{y_i}(1-\theta_i)^{n_i-y_i}$ under the normal distribution $\eta_i \sim N(\mu, \tau^2)$. THis cannot be done analytically.
(c) The top two levels of the hierarchical model are not conjugate, i.e. we cannot write $p(\eta_i|\mu, \tau, y)$ in closed form.

5.10 By (5.17): $\theta_j|\mu, \tau, y \sim N(\hat{\theta}_j, V_j)$ (see (5.17) for the defintion of $\hat{\theta}_j$ and V_j) and by (5.20) $p(\mu|\tau, y) = N(\hat{\mu}, V_\mu)$. Hence $p(\theta_j|\tau, y) = N(m, s^2)$ with

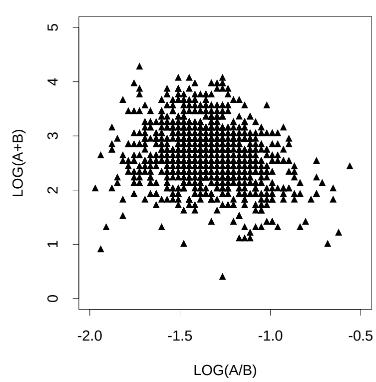
$$m = s^2(\frac{1}{\sigma_j^2}\bar{y}_{\cdot j} + \frac{1}{\tau}\hat{\mu})$$

and $1/s^2 = 1/V_j + 1/V_\mu$.[1]

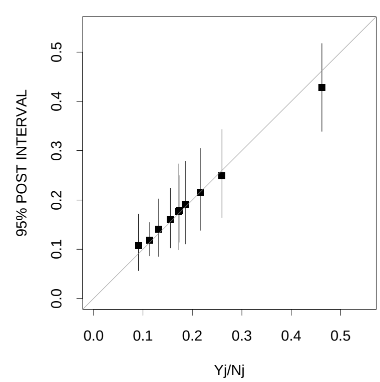
5.11 (a)–(e) See the Splus program on the 215 homepage (click "problems").
(d) The 95% posterior interval for $m = \alpha/(\alpha+\beta)$ is estimated as $[0.14, 0.30]$.
(e) The 95% posterior predictive interval for y_{11} is estimated as $[3, 50]$.



$p(\alpha, \beta|y)$



$(\alpha, \beta) \sim p(\alpha, \beta|y)$



Part (c).

[1] This follows from the general result: $X|Y \sim N(Y, \sigma^2), Y \sim N(m, \tau^2)$, then $X \sim N(m, \sigma^2+\tau^2)$.

6.9a Under H_1:

$$p(y|H_1) = \prod_{j=1}^{J} \int N(y_j|\theta_j, s_j^2) N(\theta_j|0, A^2) d\theta_j == \prod N(y_j|0, s_j^2 + A^2).$$

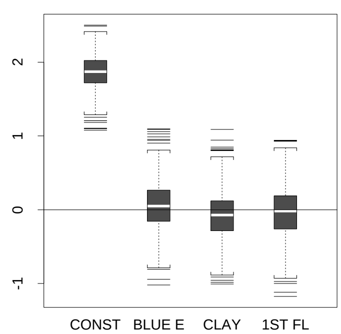
The last equality is true by Lemma 1.
Under H_2: Rewrite model H_2: $y_j \sim N(\theta, s_j^2)$, and $\theta\ N(0, A^2)$ as $y_j = \theta + z_j$, with $z_j \sim N(0, s_j^2)$, $\theta\ N(0, A^2)$, Write $x = (z_1, ...z_J, \theta)$ and $w = (y_1, ...y_J, \theta)$. Then $x \sim N(0, S)$ and $w = Ax$ with

$$A = \begin{bmatrix} 1 & 0 & .. & .. & 0 & 1 \\ 0 & 1 & 0 & .. & 0 & 1 \\ .. & .. & .. & .. & .. & .. \\ 0 & .. & .. & .. & 1 & 1 \\ 0 & .. & .. & .. & 0 & 1 \end{bmatrix}$$

Hence, by Lemma 2: $w \sim N(0, ASA')$. Find ASA' (by simple matrix multiplication). Finally, by Lemma 3 $y = (y_1...y_J) \sim N(0, T)$ with T the $J \times J$ left upper submatrix of ASA'.
Thus

$$B_{12} = \frac{\prod N(y_j|0, s_j^2 + A^2)}{N(y|0, T)}.$$

14.1 (a) I used indicators $x_{i1} = 1$ for Blue Earth county (0 otherwise); $x_{i2} = 1$ for Clay county (0 otherwise); and $x_{i3} = 1$ for first floor (0 otherwise). Let X denote the 41×4 matrix with columns $x_{i0} = 1$, $x_{i1}, \ldots, x_{i3}$ (the column of all 1's is included for an intercept). Let $y = \log(\text{radon})$. Let $\beta = (\beta_0, \ldots, \beta_3$ denote a vector of regression coefficients. We use a regression model $y = X\beta + \epsilon$, with independent normal errors $\epsilon_i \sim N(0, \sigma^2)$. With a non-informative prior $p(\beta, \sigma^2) \propto 1/\sigma^2$ I get the following posterior inference, summarized by a boxplot of the marginal posterior distributions for the parameters $\beta_0, \ldots, \beta_3$:



Posterior simulations from $p(\beta_j|y)$.

There is little posterior evidence of any significant differences, neither between Blue Earth county and Goodhue county (2nd boxplot) nor between Clay and Goodhue county. Actually, the posterior gives $Pr(\text{Blue Earth} > \text{Goodhue}) \approx 0.56$ and $Pr(\text{Clay} > \text{Goodhue}) \approx 0.40$. The largest difference is estimated for $Pr(\text{Blue Earth} > \text{Clay}) \approx 0.68$. Also, a posteriori there seems to be little difference between basement and first floor measurements (fourth boxplot).

14.3

$$p(\beta|\sigma^2, y) \quad \propto \quad \exp[-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)]$$

$$
\begin{aligned}
&= \exp[-\frac{1}{2\sigma^2}(y'y - 2y'X\beta + \beta'X'X\beta)] \\
&\propto \exp[-\frac{1}{2\sigma^2}(\beta'(X'X)\beta - 2\beta'(X'X)(X'X)^{-1}X'y)] \\
&\propto \exp[-\frac{1}{2\sigma^2}(\beta - (X'X)^{-1}X'y)'(\beta - (X'X)^{-1}X'y)] \\
&\propto N(\beta|(X'X)^{-1}X'y, \sigma^2(X'X)^{-1}) \\
&= N(\beta|\hat{\beta}, \sigma^2 V_\beta).
\end{aligned}
$$

14.7

$$
\begin{aligned}
p(\tilde{y}|\sigma^2, y) &= \int \underbrace{N(\tilde{y}|\tilde{x}'\beta, \sigma^2)}_{p(\tilde{y}|\beta,\sigma^2)} \underbrace{N(\beta|\hat{\beta}, \sigma^2 V_\beta)}_{p(\beta|\sigma^2,y)} \, d\beta \\
&= N(\tilde{x}'\beta, \sigma^2 + \sigma^2 \tilde{x}' V_\beta \tilde{x}).
\end{aligned}
$$

The last equality follows from the *Lemma*.