Bayesian Statistics in R: Assignment 4

(40 points)

1. Include a log-likelihood and deviance calculation in the JAGS code at case level (LL_{ik} and d_i in 4.1.7). The indicators r_{ik} can be obtained using the equals command

```
r[i,k] <- equals(y[i],k).
```

Run the model for 2500 iterations. Which two cases have the highest deviance? (4 points)

2. Calculate the interaction between price and catch variables and fit the classical MNL model (using mnlogit) but now including the price-catch interaction, and assuming choice varying coefficients for that interaction predictor. How far does the log-likelihood increase: by under 15 or over 15. (4 points)

3. Extend the JAGS model to have a second attribute specific predictor with choice varying effects, i.e. there is a second Z variable (see 4.1.5) formed by the interaction of 'price' and 'catch', in addition to the first Z variable which is defined simply by 'catch'. Run the model for 2500 iterations. Does the posterior mean deviance fall (a) by more than 40, or (b) by less than 40. (4 points)

4. Include extra code to (a) sample replicate data $y_{rep,i}$ (b) check whether replicate values equal actual values y_i and (c) calculate the total classification accuracy. The latter is the number of cases correctly classified divided by the sample size. Is the classification accuracy rate (a) above 25%, or (b) below 25%. (4 points)

5. There may be rater (judge) effects on the ordinal responses. We can expand the model to include a rater effect. This will involve 9 extra parameters, say omega[1:9], of which one (say omega[1]) is fixed at zero (a corner constraint). The remaining unknowns, omega[2],…omega[9], can be assigned normal priors with mean 0 and variance 1000. The linear regressor would now be

```
eta[i] <- beta[1]*contact[i] + beta[2]*temp[i]+omega[judge[i]]
```

Is the classification accuracy now (a) above 30% or (b) below 30%. (4 points)

6. Extend the model in Q5 to include an interaction between temperature and contact. Is the classification accuracy now (a) above 40% or (b) below 40%. (4 points)

7. Which predictors have posterior probability of inclusion exceeding 0.95? (2 points)

8. Modify the code so that all predictors are standardised before being include in the regression mean mu[i]. For example

z1[i] <- (complaints[i]-mean(complaints[]))/sd(complaints[])

etc.

Which predictors now have posterior probabilities of inclusion exceeding 0.75. (6 points)

9. Monitor the model selection indicators using the code developed in Q8, and extract them using jagresults, as in

```
R1 <- coda.samples(M,c("mod"),n.iter=2500)
mod.R1 <- jagsresults(R1, c("mod"))
```

Express models as 1+X2+X4 (e.g. model retaining X2 and X4 only). Which model has the highest posterior probability? (2 points)

10. Modify the code developed in Q8 to allow unknown r_j as in

$\gamma_j \sim \text{Bern}(r_j)$,

$r_j \sim \text{Beta}(1,1)$.

Which posterior mean r[j] exceed 0.5? (6 points)