Assignment -5

① a) $\text{logit}[\text{Prob}(Y_i=1)] = \sum_{j=1}^{p} x_{ij}\beta_j$

$p=7$ $\quad i=1,2\ldots 2035$

$\text{logit}(x) = \log\left(\frac{x}{1-x}\right)$

using JAGS, MCMC sampling β_j's

b) using Random effect model

$s_i = \{1,\ldots L\}$ $\quad L=65$

$\text{logit}[\text{Prob}(Y_i=1)] = \sum_{j=1}^{p} x_{ij}\beta_j + \alpha_{s_i}$

$\alpha_l \sim N(0,\tau^2)$ $\quad l=1,\ldots L$

MCMC sampling of β_j's and τ^2.

② Fitting the mixture of 3 Normal distribution.

$$Y = \theta_1 \cdot N(\mu_1,\sigma_1^2) + \theta_2 N(\mu_2,\sigma_2^2) + \theta_3 N(\mu_3,\sigma_3^2)$$

③ Given the data Y_1, N_1, Y_2, N_2

M_1: $Y_1|\lambda_1 \sim \text{Poisson}(N_1\lambda_1)$ & $Y_2|\lambda_2 \sim \text{Poisson}(N_2\lambda_2)$

M_2: $Y_1|\lambda_0 \sim \text{Poisson}(N_1\lambda_0)$ & $Y_2|\lambda_0 \sim \text{Poisson}(N_2\lambda_0)$

$\lambda_j \sim \text{Uniform}(0,c)$ $\quad \Rightarrow c=1,10$

$$\text{Bayes factor} = \frac{P(Y|M_2)}{P(Y|M_1)}$$

$$P(Y|M_2) = \iint p(y_1|\lambda_0)\, d\lambda_0 \cdot p(y_2|\lambda_0)\, d\lambda_0$$

$$= \int_0^c \int_0^c \frac{\lambda_0^{y_1} \cdot e^{-\lambda_0}}{y_1!} \cdot \frac{\lambda_0^{y_2} \cdot e^{-\lambda_0}}{y_2!}\, d\lambda_0\, d\lambda_0$$

$$= \frac{1}{y_1!\, y_2!} \int_0^c \int_0^c \lambda_0^{y_1+y_2}\, e^{-2\lambda_0}\, d\lambda_0\, d\lambda_0$$

$$P(Y|M_1) = \iint p(y_1|\lambda_1)\, p(y_2|\lambda_2)\, d\lambda_1\, d\lambda_2$$

$$= \int_0^c \int_0^c \frac{\lambda_1^{y_1} e^{-\lambda_1}}{y_1!} \cdot \frac{\lambda_2^{y_2} \cdot e^{-\lambda_2}}{y_2!}\, d\lambda_1\, d\lambda_2$$

$$= \frac{1}{y_1!\, y_2!} \int_0^c \int_0^c \lambda_1^{y_1} \cdot \lambda_2^{y_2} \cdot e^{-(\lambda_1+\lambda_2)}\, d\lambda_1\, d\lambda_2$$

$$\frac{P(Y|M_2)}{P(Y|M_1)} = \frac{\int_0^c \int_0^c \lambda_0^{y_1+y_2}\, e^{-2\lambda_0}\, d\lambda_0^2}{\int_0^c \int_0^c \lambda_1^{y_1} \lambda_2^{y_2} \cdot e^{-(\lambda_1+\lambda_2)}\, d\lambda_1\, d\lambda_2} = \text{Bayes factor.}$$

DIC :-

$\bar{D} - E[D(Y|\theta)|Y]$

$P_D = \bar{D} - D(Y|\hat{\theta})$

$DIC = \bar{D} + P_D$

WAIC

$WAIC = -2\sum_{i=1}^{n} m_i + 2P_W$

m_i = mean of $\log[f(y_i|\theta)]$

v_i = variance of $\log[f(y_i|\theta)]$

$P_W = \sum_{i=1}^{n} v_i$

(5) $Y_t|Y_{t-1} \cdots Y_1 \sim N(\beta_0 + \beta_1 Y_{t-1} + \cdots + \beta_L Y_{t-L}, \sigma^2)$

$L = 1, 2, 3, 4$

To select the best time lag L, use WAIC

$WAIC = -2\sum m_i + 2\sum v_i$

m_i & v_i are mean & variance of likelihood.