1 Data Augmentation

1.1 log of an exponential

The exponential with mean $\lambda$ has density,

$$p(t) = \frac{\exp(-t/\lambda)}{\lambda} dt$$

If we make the transformation $x = \log(t/\lambda)$ then $t = \lambda \exp(x)$ and $dt = \lambda \exp(x) dx$ so the density of $x$ is,

$$p(x) = \frac{\exp(-\lambda \exp(x)/\lambda)}{\lambda \lambda \exp(x)} dx = \exp(x - \exp(x)) dx$$

1.2 Data augmentation for Poisson regression

For a Poisson process with with expected number of events in unit time $\mu$ the intervals between successive events, $t$, follow an exponential distribution with mean $\lambda = 1/\mu$ and $x = \log(t/\lambda) = \log(\mu t) = \log(\mu) + \log(t)$, the $x$ will have the density,

$$p(x) = \exp(x - \exp(x)) dx$$

If the Poisson regression model has a linear predictor $\log(\mu) = \eta$, then $\log(t) + \eta$ will have a density $\exp(x - \exp(x))$.

Approximating $\exp(x - \exp(x))$ by a single normal distribution would be very inaccurate. So instead we use a mixture distribution based on five normal distributions $N(m_k, v_k)$ $k = 1...5$ with weights $w_k$. The probability that component $k$ generates an observed time $t$ is $w_k \phi(log(t) + \eta; m_k, v_k)dt$ where $\phi(x; m, v)$ is the density of a normal distribution with mean $m$ and variance $v$ associated with the observation $x$. The probability that the observation comes from component $k$ is therefore,

$$\frac{w_k \phi(log(t) + \eta; m_k, v_k)}{\sum_{j=1}^{5} w_j \phi(log(t) + \eta; m_j, v_j)} \quad k = 1...5$$

We can simulate the component by selecting $k$ with these probabilities.

1.3 A Gibbs Sampler

We have a model for the mean of the $i^{th}$ Poisson count that has a general form,

$$log(\mu_i) = \sum_{j=1}^{p} \beta_j x_{ij}$$

We will generate $y_i + 1$ random times between events and associate them with $y_i + 1$ different components from the 5 component mixture, $r(i, k)$, so that,

$$log(t_{ik} - log(\lambda_i)) = log(t_{ik} + log(\mu_i)) \sim N(m_{r(i,k)}, v_{r(i,k)}) \quad k = 1...y_i + 1$$
where \( v \) is the variance. Assuming that we have normal priors so that \( \beta_j \sim N(M_j, V_j) \) then the terms in the log-posterior that include \( \beta_j \) are,

\[
-\frac{1}{2} \sum_{i-1}^{n} \sum_{k=1}^{y_i+1} \left( \frac{\log(t_{ik}) + \log(\mu_i) - m_{r(i,k)}}{v_{r(i,k)}} \right)^2 - \frac{1}{2} \frac{(\beta_j - M_j)^2}{V_j}
\]

This is quadratic in \( \beta_j \) so the conditional posterior will be normally distributed. If we collect the terms in \( \beta_j^2 \) we obtain,

\[
B = -\frac{1}{2} \left[ \sum_{i-1}^{n} \sum_{k=1}^{y_i+1} \frac{x_{ij}^2}{v_{r(i,k)}} + \frac{1}{V_j} \right]
\]

and the terms in \( \beta_j \) give

\[
A = -\sum_{i-1}^{n} \sum_{k=1}^{y_i+1} \left( \frac{\log(t_{ik}) + \log(\bar{\mu}_i) - m_{r(i,k)}}{v_{r(i,k)}} \right) x_{ij} + M_j \frac{V_j}{V_j}
\]

where \( \bar{\mu}_i \) is the linear predictor without the term in \( \beta_j \). It follows that the posterior for \( \beta_j \) must be,

\[
N(A/B, \text{var} = 1/B)
\]

The updating of the precision estimate for the subject level random effect \( \tau_u \) is exactly as we had for the non-augmented analysis, that is,

\[
G \left( a_u + 29.5, \left[ \frac{1}{b_u} + \frac{1}{2} \sum_{i=1}^{59} u_i^2 \right]^{-1} \right)
\]

Similar calculations for \( \tau_e \) lead to,

\[
G \left( a_e + 118, \left[ \frac{1}{b_e} + \frac{1}{2} \sum_{i=1}^{59} \sum_{t=1}^{4} c_{it}^2 \right]^{-1} \right)
\]