MATH 523: CONNECTION BETWEEN POISSON AND MULTINOMIAL LIKELIHOODS

Binomial: If $Y_1 \sim Poisson(\mu_1)$ and $Y_2 \sim Poisson(\mu_2)$ are independent, and if $M = Y_1 + Y_2$, then we have that

$$Y_1 | M = m \sim Binomial(m, \pi)$$

where $\pi = \mu_1/(\mu_1 + \mu_2)$. Suppose we have a GLM for the two data, where $\log \mu_1 = \beta_0$, $\log \mu_2 = \beta_0 + \beta_1$. The likelihood based on data (y_1, y_2) is given by

$$L(\beta_0, \beta_1; y_1, y_2) = f_{Y_1, Y_2}(y_1, y_2; \beta_0, \beta_1)$$

so that the log-likelihood is

$$\begin{aligned} l(\beta_0, \beta_1; y_1, y_2) = \log L(\beta_0, \beta_1; y_1, y_2) &= y_1 \log \mu_1 - \mu_1 + y_2 \log \mu_2 - \mu_2 + \text{const.} \\ &= y_1 \beta_0 - e^{\beta_0} + y_2(\beta_0 + \beta_1) - e^{\beta_0 + \beta_1} + \text{const.} \\ &= (y_1 + y_2)\beta_0 + y_2 \beta_1 - e^{\beta_0}(1 + e^{\beta_1}) + \text{const.} \end{aligned}$$

Reparameterizing to use $\psi = \mu_1 + \mu_2 = e^{\beta_0} + e^{\beta_0 + \beta_1} = e^{\beta_0}(1 + e^{\beta_1})$, we may re-write the likelihood for (ψ, β_1) as

$$l(\psi, \beta_1; y_1, y_2) = (y_1 + y_2) \log \psi - \psi + y_2 \beta_1 - (y_1 + y_2) \log(1 + e^{\beta_1}) + \text{const.}$$

and, equivalently, if $m = y_1 + y_2$,

$$\begin{aligned} l(\psi, \beta_1; m, y_2) &= m \log \psi - \psi + y_2 \beta_1 - m \log(1 + e^{\beta_1}) + \text{const.} \\ &= l_M(\psi; m) + l_{Y_2|M}(\beta_1; y_2, m) + \text{const.} \end{aligned}$$

where

$$l_M(\psi; m) = m \log \psi - \psi \qquad l_{Y_2|M}(\beta_1; y_2, m) = y_2 \beta_1 - m \log(1 + e^{\beta_1})$$

The first term $l_M(\psi; m)$ is (up to an additive constant) the log-likelihood for a Poisson model with mean $\psi = e^{\beta_0}(1 + e^{\beta_1}) = \mu_1 + \mu_2$. The second term is (up to an additive constant) the log-likelihood for a Binomial model with parameters m and

$$\pi = \frac{e^{\beta_1}}{1 + e^{\beta_1}} = \frac{e^{\beta_0 + \beta_1}}{e^{\beta_0} + e^{\beta_0 + \beta_1}} = \frac{\mu_2}{\mu_1 + \mu_2}.$$

Hence inference about β_1 can be made via the **conditional** likelihood for Y_2, given $M = Y_1 + Y_2$ is observed to be m; this follows as, in the new parameterization, terms involving β_1 appear only in the conditional likelihood.

Multinomial I: This result may be extended to the K random variable case, and a multinomial model. Suppose that $Y_k \sim Poisson(\mu_k)$ for $k = 1, \ldots, K$ are independent, and

$$\log \mu_k = \beta_0 + \beta_1 x_k$$

say (that is, retaining a two-parameter model for the moment). The calculation above goes through as before with

$$\psi = \sum_{k=1}^{K} e^{\beta_0 + \beta_1 x_k} = e^{\beta_0} \sum_{k=1}^{K} e^{\beta_1 x_k}$$

and $M = Y_1 + \cdots + Y_K$. Then if $m = y_1 + \cdots + y_K$,

$$\begin{aligned} l(\psi, \beta_1; m, y_1, \ldots, y_{K-1}) &= m \log \psi - \beta_1 \sum_{k=1}^{K} x_k y_k - m \log \left(\sum_{k=1}^{K} e^{\beta_1 x_k} \right) + \text{const.} \\ &= l_M(\psi; m) + l_{Y|M}(\beta_1; y_1, \ldots, y_K, m) + \text{const.} \end{aligned}$$

The first term $l_M(\psi; m)$ is (up to an additive constant) the log-likelihood for a Poisson model with mean $\psi = \mu_1 + \cdots + \mu_K$. The second term is (up to an additive constant) the log-likelihood for a Multinomial model with parameters m and $(\pi_1, \ldots, \pi_K)$ where

$$\pi_k = \frac{e^{\beta_1 x_k}}{\sum_{l=1}^{K} e^{\beta_1 x_l}} = \frac{e^{\beta_0 + \beta_1 x_k}}{\sum_{l=1}^{K} e^{\beta_0 + \beta_1 x_l}} = \frac{\mu_k}{\mu_1 + \cdots + \mu_K}.$$

We can see this result as for the multinomial case

$$L(\pi_1, \ldots, \pi_K; m, y_1, \ldots, y_K) = \frac{m!}{y_1! \ldots y_K!} \pi_1^{y_1} \ldots \pi_K^{y_K}$$

yielding a log-likelihood

$$\begin{aligned} l(\pi_1, \ldots, \pi_K; m, y_1, \ldots, y_K) &= \sum_{k=1}^{K} y_k \log \pi_k + \text{const.} = \sum_{k=1}^{K} y_k \log \left(\frac{e^{\beta_1 x_k}}{\sum_{l=1}^{K} e^{\beta_1 x_l}} \right) + \text{const.} \\ &= \sum_{k=1}^{K} y_k \left(\beta_1 x_k - \log \left(\sum_{l=1}^{K} e^{\beta_1 x_l} \right) \right) + \text{const.} \\ &= \beta_1 \sum_{k=1}^{K} y_k x_k - m \log \left(\sum_{l=1}^{K} e^{\beta_1 x_l} \right) + \text{const.} \end{aligned}$$

as required. In the above calculations we may write $\beta_k = \beta_1 x_k$ for convenience.

Multinomial II: Finally, suppose that we have a Poisson model where $\log \mu_k = \beta_0 + \beta_k$, with $\beta_1 = 0$ for identifiability. Writing

$$\psi = \sum_{k=1}^{K} \mu_k = \sum_{k=1}^{K} e^{\beta_0 + \beta_k} = e^{\beta_0} \left(1 + \sum_{k=2}^{K} e^{\beta_k} \right)$$

and conditioning on $M = Y_1 + \cdots + Y_K$ taking the value m, following through the identical steps, we have the same factorization

$$l(\psi, \beta_2, \ldots, \beta_K; m, y_1, \ldots, y_{K-1}) = l_M(\psi; m) + l_{Y|M}(\beta_2, \ldots, \beta_K; y_1, \ldots, y_K, m) + \text{const.}$$

where

$$\begin{aligned} l_M(\psi; m) &= m \log \psi - \psi \\ l_{Y|M}(\beta_2, \ldots, \beta_K; y_1, \ldots, y_K, m) &= \sum_{k=1}^{K} y_k \beta_k - m \log \left(\sum_{l=1}^{K} e^{\beta_l} \right). \end{aligned}$$

The conditional distribution that yields the log-likelihood $l_{Y|M}(\beta_2, \ldots, \beta_K; y_1, \ldots, y_K, m)$ is a multinomial distribution with probabilities

$$\pi_k = \frac{e^{\beta_k}}{\sum_{l=1}^{K} e^{\beta_l}} = \frac{e^{\beta_0 + \beta_k}}{\sum_{l=1}^{K} e^{\beta_0 + \beta_l}} = \frac{\mu_k}{\mu_1 + \cdots + \mu_K}.$$

Note that as $\beta_1 = 0$,

$$\pi_1 = \frac{1}{1 + \sum_{l=2}^{K} e^{\beta_l}} \qquad \pi_k = \frac{e^{\beta_k}}{1 + \sum_{l=2}^{K} e^{\beta_l}}, \quad k = 2, \ldots, K.$$