Negative Binomial Models

The negative binomial distribution can be expressed with two parameters r and b with mean rb and variance rb(1+b). The log of the probability of k events is:

$$\log(p_k) = \sum_{i=0}^{k-1} \log(r+i) + k\log(b) - \log(k!) - (r+k)\log(1+b)$$

When the mean is expressed as a function of the parameters of a model, r can be replaced by m/b, or b can be replaced by m/r. These give two somewhat different effects. In the first case, with parameters m and b, the mean is m and the variance is m(1+b), and so is proportional to the mean. In the second case, with parameters m and r, the mean is still m, but now the variance is m(1+m/r), which is quadratic in the mean. The latter form is a member of the exponential family, and is used in GLM. The former is more like a Poisson with over-dispersion, with the variance a multiple of the mean.

For MLE, the derivative of the log probability with respect to m is needed. In the m,b case where r is replaced by m/b this is:

$$-\frac{\log(1+b)}{b} + \sum_{i=0}^{k-1} \frac{1}{m+bi}$$

For the m,r GLM case where b is replaced by m/r it is:

$$\frac{k}{m} - \frac{r+k}{r+m}$$

In the mortality study, $k = D_{w,d}$ is the number of deaths in the w,d cell and $m = m_{w,d}E_{w,d}$. Then in the m,b case:

$$\partial L_{w,d} / \partial m_{w,d} = -E_{w,d}\frac{\log(1+b)}{b} + \sum_{i=0}^{D_{w,d}-1} \frac{1}{m_{w,d}E_{w,d}+bi}$$

In the m,r case:

$$\partial L_{w,d} / \partial m_{w,d} = \frac{D_{w,d}}{m} - \frac{r+D_{w,d}}{r/E_{w,d}+m}$$