MATH 680 - EXERCISES 1

The **negative binomial** model is a useful model for count data. The mass function is

$$f(x|\theta,\alpha) = \frac{\Gamma(x+\alpha)}{\Gamma(x+1)\Gamma(\alpha)}\theta^{\alpha}(1-\theta)^{x} \qquad x = 0,1,2,\ldots.$$

where $\alpha > 0$, and $0 < \theta < 1$. In this model

$$\mathbb{E}[X] = \alpha\frac{1-\theta}{\theta} = \mu \qquad\qquad \mathbb{V}ar[X] = \alpha\frac{(1-\theta)}{\theta^2}$$

and this suggests an alternative parameterization in terms if μ. We have

$$\theta = \frac{\alpha}{\mu+\alpha}$$

and thus

$$f(x|\mu,\alpha) = \frac{\Gamma(x+\alpha)}{\Gamma(x+1)\Gamma(\alpha)}\frac{\mu^{x}\alpha^{\alpha}}{(\mu+\alpha)^{\alpha+x}} \qquad x = 0,1,2,\ldots.$$

In this parameterization

$$\mathbb{V}ar[X] = \mu + \frac{\mu^2}{\alpha}.$$

An alternative motivation for this model is as a Gamma mixture of Poisson random variables, that is, it is the marginal model for X if

$$X|V \sim \text{Poisson}(\mu V) \qquad V \sim \text{Gamma}(\alpha,\alpha)$$

(a) Using one of the methods described in lectures, compute the ML estimates for (θ,α) for the count data in the following example: counts of the number of red mites on 150 apple leaves were recorded

Number of mites per leaf	0	1	2	3	4	5	6	7	8+
Frequency	70	38	17	10	9	3	1	1	0

Note: Make sure to construct the likelihood correctly.

Reference: Fitting the Negative Binomial Distribution to Biological Data, C. I. Bliss and R. A. Fisher, *Biometrics* Vol. 9, No. 2 (Jun., 1953), pp. 176-200.

(b) Check your answer from (a) using the pre-written functions in `R`.

(c) *For those students who have taken MATH 523 or equivalent.*

Suppose that the negative binomial distribution is to be used as the basis of a GLM. Suggest suitable link functions for this Exponential-Dispersion Family model.

The `MASS` library in `R` contains a data set `quine`, with a count data response, and four factor predictors. The library also contains a function, `glm.nb`, that can be used to obtain ML estimates. See for example

`http://stat.ethz.ch/R-manual/R-devel/library/MASS/html/glm.nb.html`

After establishing that the Poisson model is not sufficient for these data, proceed to fit the negative binomial model and examine the resulting estimates and their standard errors.