SMM637 Exercises - GLM

Hurn *et al* (1945) reported an experiment which investigated the relationship between normal blood plasma concentration in blood and clotting time. Samples of normal blood plasma were diluted to nine different percentage concentrations, using plasma free of the clotting factor prothrombin. Clotting was then induced in each of the samples using two different lots of thromboplastin, and the clotting time in seconds was recorded.

The data were entered into R as three vectors: a variable `conc` giving the concentration of normal plasma (as a percentage), a factor `lot` at two levels (1 and 2) denoting the lot of thromboplastin, and a variable `clottime` giving the clotting time in seconds. Treating `clottime` was treated as the response variable.

An appropriate model for these data was considered to be a gamma regression with a reciprocal link function (the canonical link). To obtain a good fit it was necessary to transform `conc` by taking logs (to give `lconc`). Within this framework, the following model was fitted.

```
> lconc <- log(conc)

> modC.glm <- glm(clottime ~ lconc + lot + lconc:lot, family = Gamma)
> summary(modC.glm)

Deviance Residuals:
         Min           1Q       Median         3Q        Max
 -0.05573777 -0.03547972 -0.008216152 0.0260727 0.08641118

Coefficients:
                    Value   Std. Error    t value
(Intercept) -0.016554382 0.0008654840 -19.127311
      lconc  0.015343115 0.0003871945  39.626377
        lot -0.007354088 0.0016779383  -4.382812
  lconc:lot  0.008256099 0.0007352771  11.228555

(Dispersion Parameter for Gamma family taken to be 0.0021297 )
    Null Deviance: 7.708668 on 17 degrees of freedom
Residual Deviance: 0.0294015 on 14 degrees of freedom
```

1. What model is being fitted?
2. The experimenters wanted to know whether the relationship between normal plasma concentration and clotting time differed between the two lots of thromboplastin.
3. We can now write down the fitted equations for the mean clotting time, separately for Lot 1 and Lot 2. The fitted model can be used to predict the mean clotting time in different circumstances. For example, what would be the mean clotting time for a normal plasma concentration of 50% using thromboplastin from Lot 2?
4. On the basis of the outputs given, would we expect that exponential regression would provide a suitable model for these data?