MATH523 Assignment 2

Kabilan Sriranjan

February 20, 2018

A5

a)

The expected Fisher information for a Gamma GLM is given by

$$I(\beta) = X^T W X$$

$$= \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix} diag\left(\frac{1}{a(\phi_i)b''(\theta_i)g'(\theta_i)}\right) \begin{bmatrix} 1 & x_1 \\ \cdots & \cdots \\ 1 & x_n \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{a(\phi_1)b''(\theta_1)g'(\theta_1)} & \cdots & \frac{1}{a(\phi_n)b''(\theta_n)g'(\theta_n)^2} \\ \frac{x_1}{a(\phi_1)b''(\theta_1)g'(\theta_1)} & \cdots & \frac{x_n}{a(\phi_n)b''(\theta_n)g'(\theta_n)^2} \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ \cdots & \cdots \\ 1 & x_n \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^n \frac{1}{a(\phi_i)b''(\theta_i)g'(\theta_i)} & \sum_{i=1}^n \frac{x_i}{a(\phi_i)b''(\theta_i)g'(\theta_i)} \\ \sum_{i=1}^n \frac{x_i}{a(\phi_i)b''(\theta_i)g'(\theta_i)} & \sum_{i=1}^n \frac{x_i^2}{a(\phi_i)b''(\theta_i)g'(\theta_i)^2} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^n \frac{1}{\frac{\mu_i^2}{\nu_i}g'(\theta_i)^2} & \sum_{i=1}^n \frac{x_i}{\frac{\mu_i^2}{\nu_i}g'(\theta_i)^2} \\ \sum_{i=1}^n \frac{x_i}{\frac{\mu_i^2}{\nu_i}g'(\theta_i)^2} & \sum_{i=1}^n \frac{x_i^2}{\frac{\mu_i^2}{\nu_i}g'(\theta_i)^2} \end{bmatrix}$$

In the case of the canonical link we can simplify this to

$$I(\beta) = \begin{bmatrix} \sum_{i=1}^n \mu_i^2 \nu_i & \sum_{i=1}^n \mu_i^2 x_i \nu_i \\ \sum_{i=1}^n \mu_i^2 x_i \nu_i & \sum_{i=1}^n \mu_i^2 x_i^2 \nu_i \end{bmatrix}$$

b)

The asymptotic variances of $\widehat{\beta}_0$ and $\widehat{\beta}_1$ are given by the reciprocal of the diagonal elements of the expected Fisher information

$$var(\beta_0) = \frac{1}{\sum_{i=1}^n \nu_i}$$

$$var(\beta_1) = \frac{1}{\sum_{i=1}^n x_i^2 \nu_i}$$

c)

d)

e)

$$D(y, \hat{\mu}) = 2 \sum_{i=1}^{n} w_i (y_i (\bar{\theta}_i - \hat{\theta}_i) - b(\bar{\theta}_i) + b(\hat{\theta}_i))$$

$$= \sum_{i=1}^{n} w_i (-\frac{1}{y_i} + \frac{1}{\hat{\mu}_i}) - log(y_i) + log(\hat{\mu}_i))$$

$$= 2 \sum_{i=1}^{n} w_i (\frac{y_i - \hat{\mu}_i}{\hat{\mu}_i} - log(\frac{y_i}{\hat{\mu}_i}))$$

f)

Pearson residuals

$$\frac{y_i - \hat{\mu}_i}{\hat{\mu}} \sqrt{w_i}$$

Anscombe residuals

$$\frac{3(y_i^{\frac{1}{3}} - \hat{\mu}_i^{\frac{1}{3}})}{\hat{\mu}_i^{\frac{1}{6}}}$$

Deviance residuals

$$sign(y_i - \hat{\mu}_i) \sqrt{2w_i (\frac{y_i - \hat{\mu}_i}{\hat{\mu}_i} - log(\frac{y_i}{\hat{\mu}_i}))}$$

A6

For a Normal GLM we have $\theta_i = \mu_i$ and $b(\theta) = \frac{\theta^2}{2}$. The deviance is given by

$$D(y, \hat{\mu}) = 2 \sum_{i=1}^{n} (y_i (y_i - \hat{\mu}_i) - \frac{y_i^2}{2} + \frac{\hat{\mu}_i^2}{2})$$

$$= \sum_{i=1}^{n} (y_i - \hat{\mu}_i)^2$$

So the deviance is just the sum of the squared residuals In a linear regression we may consider the quantity $SS_{res}(\beta_2|\beta_1)$ to compare the residual sum of squares of a model of just β_1 with one where we include β_2. This is exactly the difference of deviances.

A7

a)

```r
library(glmbb)
data(crabs)

fullModel <- glm(satell~(weight+color)^2, data=crabs, family=poisson)
summary(fullModel)
```

```
##
## Call:
## glm(formula = satell ~ (weight + color)^2, family = poisson,
##     data = crabs)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.9412  -1.8457  -0.6278   0.8346   4.7689
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)        -1.463e+00  4.851e-01  -3.016  0.00256 **
## weight              9.283e-04  1.839e-04   5.048 4.47e-07 ***
## colordarker        -5.214e-02  9.050e-01  -0.058  0.95406
## colorlight          3.408e+00  1.127e+00   3.023  0.00250 **
## colormedium         1.386e+00  5.345e-01   2.594  0.00949 **
## weight:colordarker  5.232e-05  3.645e-04   0.144  0.88587
## weight:colorlight  -1.134e-03  4.287e-04  -2.645  0.00818 **
## weight:colormedium -4.466e-04  2.002e-04  -2.231  0.02570 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 632.79  on 172  degrees of freedom
## Residual deviance: 541.08  on 165  degrees of freedom
## AIC: 912.38
##
## Number of Fisher Scoring iterations: 6
```

This is the summary of the model with weight, colour, and interaction terms between weight and colour. I used the poisson GLM with the canonical link because the number of satellites could theoretically be arbitarily large so Binomial wouldn't make sense.

b)

```r
X <- model.matrix(fullModel)
W <- fullModel$weights

beta <- fullModel$coefficients
beta2 <- beta[6:8]

I <- t(X) %*% diag(W) %*% X
I2 <- I[6:8,6:8]

waldTest <- t(beta2) %*% I2 %*% beta2
```

```r
pchisq(waldTest, 3, lower.tail=FALSE)
```

```
##              [,1]
## [1,] 7.393929e-203
```

The Wald test is based off the result that

$$\widehat{\beta}^{(1)T} X^T W X \widehat{\beta}^{(1)} \approx \chi^2_{p_1}$$

where X is the design matrix but only for the predictors we are testing, $\beta^{(1)}$ are the parameters for the predictors were are testing, and p_1 is the nunmber of predictors we are testing.

c)

```r
reducedModel <- glm(satell~weight+color, data=crabs, family=poisson)
devTest <- deviance(reducedModel) - deviance(fullModel)
pchisq(devTest, 3, lower.tail=FALSE)
```

```
## [1] 0.01331655
```

When comparing the full model to a model without interaction terms using the chi-squared test we get a low p-value. This suggests that we are justified in using the more complicated model.

d)

```r
newColor <- crabs$color
levels(newColor) <- c("dark", "dark", "light", "medium")

model2 <- glm(satell~(weight+newColor)^2, data=crabs, family=poisson)
summary(model2)
```

```
##
## Call:
## glm(formula = satell ~ (weight + newColor)^2, family = poisson,
##     data = crabs)
##
## Deviance Residuals:
##     Min      1Q   Median      3Q      Max
## -2.9412  -1.8470  -0.6278   0.8346   4.7574
##
## Coefficients:
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)            -1.4451144  0.4013956  -3.600 0.000318 ***
## weight                  0.0009299  0.0001559   5.963 2.48e-09 ***
## newColorlight           3.3893600  1.0936607   3.099 0.001941 **
## newColormedium          1.3682995  0.4598657   2.975 0.002926 **
## weight:newColorlight   -0.0011354  0.0004175  -2.720 0.006536 **
## weight:newColormedium  -0.0004482  0.0001749  -2.563 0.010380 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
```

```
##     Null deviance: 632.79  on 172  degrees of freedom
## Residual deviance: 541.27  on 167  degrees of freedom
## AIC: 908.57
##
## Number of Fisher Scoring iterations: 6
```

Looking at the summary of the original model we see that the parameter for the "darker" level of our factor predictor is insignicifcant. This is possibly due to darker not being very different from dark. Hence I merged all the dark and darker observations to be the same thing.

e)

```r
res_p <- residuals(model2, "pearson")
plot(fitted(model2), res_p, xlab="Satellites", ylab="Pearson residual")
title("Pearson Residuals vs Fitted Response")
```

Pearson Residuals vs Fitted Response
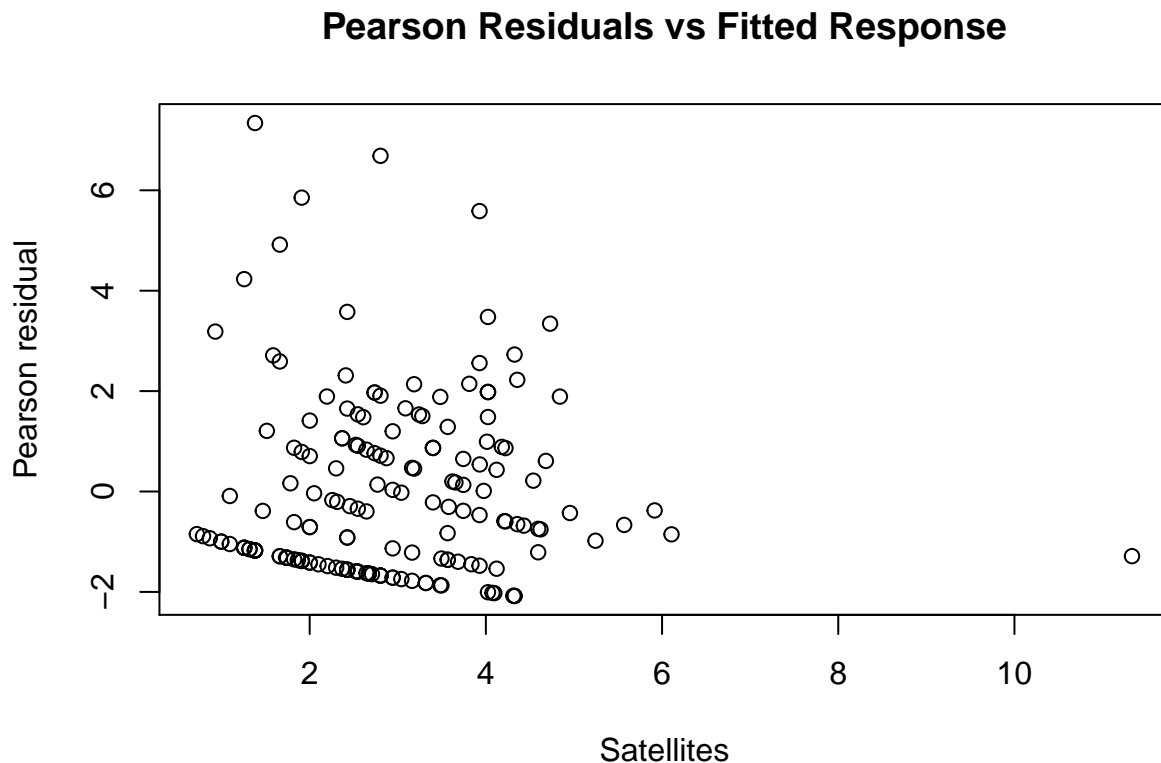

```r
res_d <- residuals(model2, "deviance")
plot(fitted(model2), res_d, xlab="Satellites", ylab="Deviance residual")
title("Deviance Residuals vs Fitted Response")
```

Deviance Residuals vs Fitted Response
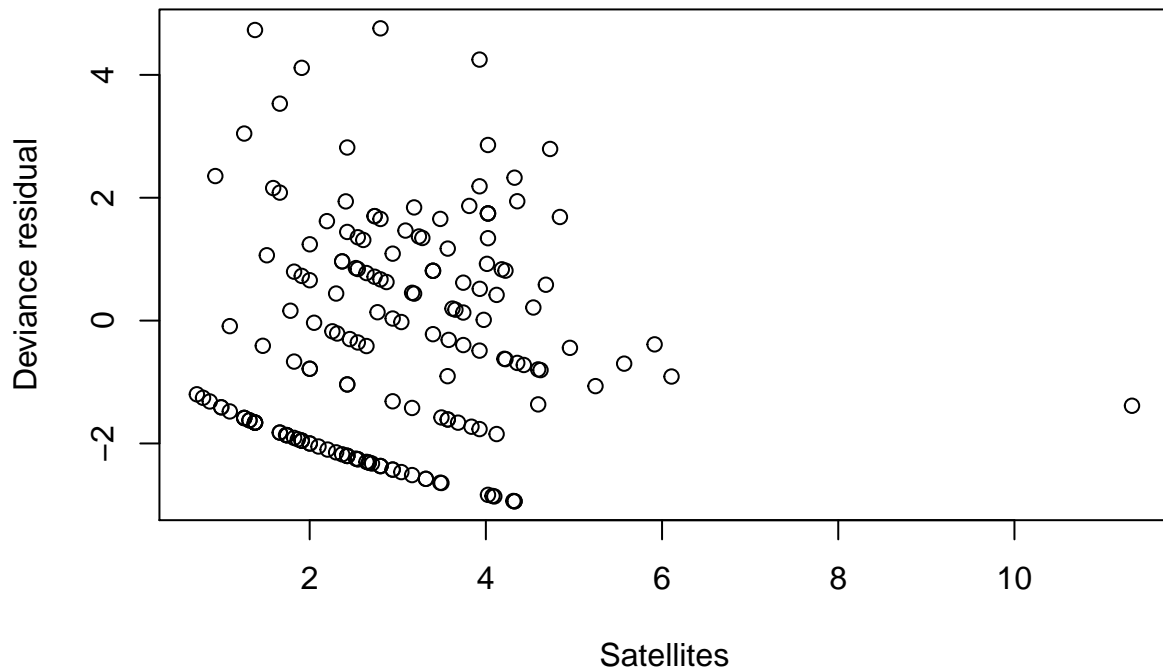

```
deviance(model2)
```

```
## [1] 541.2675
```

The residual plots appear to have some patterns but this is likely due to the discreteness of the response variable. Other than the lines they appear to be a null plot, which is good. Also the deviance is almost equal to the deviance of the original model, so we may as well take the simpler model as we aren't getting a worse deviance by doing so.