Homework 5

Yuki Joyama

2024-03-18

```r
# data prep
df_crab = read.table("/Users/yukijoyama/Library/CloudStorage/GoogleDrive-jikeyu1995@gmail.com/My Drive/
df_para = read.table("/Users/yukijoyama/Library/CloudStorage/GoogleDrive-jikeyu1995@gmail.com/My Drive/
```

1

(a)

I will fit a Poisson model (M1) with log link with carapace width (W) as the single predictor.

```r
# M1: Poisson model with log link
m1_fit <- glm(Sa ~ W, family = poisson(link = "log"), data = df_crab)
summary(m1_fit)
```

```
##
## Call:
## glm(formula = Sa ~ W, family = poisson(link = "log"), data = df_crab)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.30476    0.54224  -6.095  1.1e-09 ***
## W            0.16405    0.01997   8.216  < 2e-16 ***
## ---
## 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: 567.88  on 171  degrees of freedom
## AIC: 927.18
##
## Number of Fisher Scoring iterations: 6
```

```r
exp(m1_fit$coefficients)
```

```
## (Intercept)           W
##  0.03670812  1.17826744
```

The expected number of satellites (Sa) given carapace width (W) = 0 is 0.037. With every unit increase in W, the expected number of Sa has multiplicative effect of 1.178 on $\mu = E(Y)$.

```
# Goodness of fit
G = sum(residuals(m1_fit, type = "pearson") ^ 2)
G
```

```
## [1] 544.157
```

```
1 - pchisq(G, 2)
```

```
## [1] 0
```

Thus, we reject the null. The model does not have a good fit.

(b)

Now, I will fit a Poisson model (M2) with log link with carapace width (W) and weight (Wt) as predictors.

```
# M2: W and Wt as predictors
m2_fit <- glm(Sa ~ W + Wt, family = poisson(link = "log"), data = df_crab)
summary(m2_fit)
```

```
##
## Call:
## glm(formula = Sa ~ W + Wt, family = poisson(link = "log"), data = df_crab)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.29168    0.89929  -1.436  0.15091
## W            0.04590    0.04677   0.981  0.32640
## Wt           0.44744    0.15864   2.820  0.00479 **
## ---
## 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: 559.89  on 170  degrees of freedom
## AIC: 921.18
##
## Number of Fisher Scoring iterations: 6
```

```
exp(m2_fit$coefficients)
```

```
## (Intercept)           W          Wt
##    0.274809    1.046968    1.564296
```

```
# compare with M1
anova(m1_fit, m2_fit)
```

```
## Analysis of Deviance Table
##
## Model 1: Sa ~ W
## Model 2: Sa ~ W + Wt
##   Resid. Df Resid. Dev Df Deviance
## 1       171     567.88
## 2       170     559.89  1   7.9934
```

(c)

```
# over-dispersion in M2
```

```
# adjust for over-dispersion
```

2

a

b

c