# Lab Assignment 11

Kaushik Raj V Nadar (208160499)

2024-10-29

## Problem 1

Consider the Aircraft Damage Data in which the response variable is the number of locations where damage was inflicted on the aircraft. The regressor x1 is an indicator variable (type of aircraft), and the other regressors x2 and x3 are bomb load (in tons) and total months of aircrew experience. Fit Poisson regression model and test the hypothesis for the significance of regressors.

```r
# Create the dataset
aircraft_data <- data.frame(
  y = c(0,1,0,0,0,0,1,0,0,2,1,1,1,1,2,3,1,1,1,2,0,1,1,2,
        5,1,1,5,5,7),
  x1 = c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1,
         1,1,1,1,1,1),
  x2 = c(4,4,4,5,5,5,6,6,6,7,7,7,8,8,8,7,7,7,10,10,10,12,
         12,12,8,8,8,14,14,14),
  x3 = c(91.5,84.0,76.5,69.0,61.5,80.0,72.5,65.0,57.5,
         50.0,103.0,95.5,88.0,80.5,73.0,
         116.1,100.6,85.0,69.4,53.9,112.3,96.7,81.1,
         65.6,50.0,120.0,104.4,88.9,73.7,57.8)
)

# Fit Poisson regression model
poisson_model <- glm(y ~ x1 + x2 + x3,
                     family = poisson(link = "log"),
                     data = aircraft_data)

# Display model summary
summary(poisson_model)
```

```
Call:
glm(formula = y ~ x1 + x2 + x3, family = poisson(link = "log"),
    data = aircraft_data)

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.406023   0.877489  -0.463   0.6436
x1           0.568772   0.504372   1.128   0.2595
x2           0.165425   0.067541   2.449   0.0143 *
x3          -0.013522   0.008281  -1.633   0.1025
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 53.883  on 29  degrees of freedom
Residual deviance: 25.953  on 26  degrees of freedom
AIC: 87.649

Number of Fisher Scoring iterations: 5
```

```r
# Calculate pseudo R-squared (1 - residual deviance/null deviance)
pseudo_r2 <- 1 - poisson_model$deviance/poisson_model$null.deviance

# Perform analysis of deviance
anova_result <- anova(poisson_model, test = "Chisq")

# Print additional model statistics
cat("\nModel Statistics:\n")
```

```
Model Statistics:
```

```r
cat("Null Deviance:", poisson_model$null.deviance,
    "on", poisson_model$df.null, "degrees of freedom\n")
```

```
Null Deviance: 53.88306 on 29 degrees of freedom
```

```r
cat("Residual Deviance:", poisson_model$deviance,
    "on", poisson_model$df.residual, "degrees of freedom\n")
```

```
Residual Deviance: 25.95316 on 26 degrees of freedom
```

```r
cat("AIC:", poisson_model$aic, "\n")
```

```
AIC: 87.64922
```

```r
cat("Pseudo R-squared:", pseudo_r2, "\n")
```

```
Pseudo R-squared: 0.5183429
```

```r
# Display ANOVA results
cat("\nAnalysis of Deviance Table:\n")
```

```

Analysis of Deviance Table:
```

```r
print(anova_result)
```

```
Analysis of Deviance Table

Model: poisson, link: log

Response: y

Terms added sequentially (first to last)

     Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
NULL                    29     53.883
x1    1  15.5996        28     38.283 7.827e-05 ***
x2    1   9.6492        27     28.634  0.001894 **
x3    1   2.6811        26     25.953  0.101543
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Calculate and print odds ratios and confidence intervals
coef_exp <- exp(coef(poisson_model))
conf_int <- exp(confint(poisson_model))
```

```
Waiting for profiling to be done...
```

```r
coef_table <- cbind(Estimate = coef_exp,
                    CI_lower = conf_int[,1],
                    CI_upper = conf_int[,2])

cat("\nExponentiated Coefficients (Rate Ratios) with 95% CI:\n")
```

```
Exponentiated Coefficients (Rate Ratios) with 95% CI:
```

```r
print(coef_table)
```

```
             Estimate  CI_lower CI_upper
(Intercept) 0.6662950 0.1109979 3.504006
x1          1.7660977 0.6529760 4.792330
x2          1.1798949 1.0358915 1.351870
x3          0.9865687 0.9704577 1.002674
```

```r
# Perform Walt's test for analysis of deviance
library(lmtest)
```

```
Loading required package: zoo


Attaching package: 'zoo'


The following objects are masked from 'package:base':

    as.Date, as.Date.numeric
```

```r
walt_result <- waldtest(poisson_model)
cat("\nWald's Test Results:\n")
```

```
Wald's Test Results:
```

```r
print(walt_result)
```

```
Wald test

Model 1: y ~ x1 + x2 + x3
Model 2: y ~ 1
  Res.Df Df      F    Pr(>F)
1     26
2     29 -3 8.7541 0.0003508 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Problem 2

Following Table contains data from an experiment conducted to investigate the three factors x1 = length, x2 = amplitude, and x3 = load on the cycles to failure y of worsted yarn. The regressor variables are coded, and readers who have familiarity with designed experiments will recognize that the experimenters here used a 33 factorial design. Fit a GLM using the gamma distribution and the log link.

```r
# Create the dataset
yarn_data <- data.frame(
  y = c(674, 1414, 3636, 338, 1022, 1568, 170, 442, 1140, 370, 1198,
        3184, 266, 620, 1070, 118, 332, 884, 292, 634, 2000, 210,
        438, 566, 90, 220, 360),
  x1 = c(-1, 0, 1, -1, 0, 1, -1, 0, 1, -1, 0, 1, -1, 0, 1, -1, 0, 1,
         -1, 0, 1, -1, 0, 1, -1, 0, 1),
  x2 = c(-1, -1, -1, 0, 0, 0, 1, 1, 1, -1, -1, -1, 0, 0, 0, 1, 1, 1,
         -1, -1, -1, 0, 0, 0, 1, 1, 1),
  x3 = c(-1, -1, -1, -1, -1, -1, -1, -1, -1, 0, 0, 0, 0, 0, 0, 0, 0,
         0, 1, 1, 1, 1, 1, 1, 1, 1, 1)
)

# Fit Gamma regression model with log link
gamma_model <- glm(y ~ x1 + x2 + x3, family = Gamma(link = "log"), data = yarn_data)

# Display model summary
summary(gamma_model)
```

```
Call:
glm(formula = y ~ x1 + x2 + x3, family = Gamma(link = "log"),
    data = yarn_data)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.34891    0.03422 185.511  < 2e-16 ***
x1           0.84251    0.04192  20.100 4.34e-16 ***
x2          -0.63132    0.04192 -15.062 2.10e-13 ***
x3          -0.38513    0.04192  -9.188 3.68e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 0.0316243)

    Null deviance: 22.88613  on 26  degrees of freedom
Residual deviance:  0.76939  on 23  degrees of freedom
AIC: 332.76

Number of Fisher Scoring iterations: 5
```

```r
# Calculate pseudo R-squared (1 - residual deviance/null deviance)
pseudo_r2 <- 1 - gamma_model$deviance/gamma_model$null.deviance

# Perform analysis of deviance
anova_result <- anova(gamma_model, test = "Chisq")

# Print additional model statistics
cat("\nModel Statistics:\n")
```

```
Model Statistics:
```

```r
cat("Null Deviance:", gamma_model$null.deviance,
    "on", gamma_model$df.null, "degrees of freedom\n")
```

```
Null Deviance: 22.88613 on 26 degrees of freedom
```

```r
cat("Residual Deviance:", gamma_model$deviance,
    "on", gamma_model$df.residual, "degrees of freedom\n")
```

```
Residual Deviance: 0.7693924 on 23 degrees of freedom
```

```r
cat("AIC:", gamma_model$aic, "\n")
```

```
AIC: 332.757
```

```r
cat("Pseudo R-squared:", pseudo_r2, "\n")
```

```
Pseudo R-squared: 0.9663817
```

```r
# Display ANOVA results
cat("\nAnalysis of Deviance Table:\n")
```

```
Analysis of Deviance Table:
```

```r
print(anova_result)
```

```
Analysis of Deviance Table

Model: Gamma, link: log

Response: y

Terms added sequentially (first to last)

     Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
NULL                    26    22.8861
x1    1  12.6758        25    10.2104 < 2.2e-16 ***
x2    1   6.8645        24     3.3459 < 2.2e-16 ***
x3    1   2.5765        23     0.7694 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Calculate and print odds ratios and confidence intervals
coef_exp <- exp(coef(gamma_model))
conf_int <- exp(confint(gamma_model))
```

```
Waiting for profiling to be done...
```

```r
coef_table <- cbind(Estimate = coef_exp,
                    CI_lower = conf_int[,1],
                    CI_upper = conf_int[,2])

cat("\nExponentiated Coefficients (Rate Ratios) with 95% CI:\n")
```

```
Exponentiated Coefficients (Rate Ratios) with 95% CI:
```

```r
print(coef_table)
```

```
               Estimate    CI_lower    CI_upper
(Intercept) 571.8703665 535.1664797 612.0105689
x1            2.3221878   2.1365750   2.5239201
x2            0.5318889   0.4899942   0.5773708
x3            0.6803594   0.6260482   0.7393754
```

```r
walt_result <- waldtest(poisson_model)
cat("\nWald's test results:\n")
```

```
Wald's test results:
```

```r
print(walt_result)
```

```
Wald test

Model 1: y ~ x1 + x2 + x3
Model 2: y ~ 1
  Res.Df Df      F    Pr(>F)
1     26
2     29 -3 8.7541 0.0003508 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```