TMA4315: Compulsory exercise 2 Logistic regression and Poisson regression

Group XX: Henrik Syversveen Lie, Mikal Stapnes, Oliver Byhring

10.10.2018

Contents

Part 1: Logistic regression

a)

We let y_i be the number of successfull ascents, and n_i be the total number of attempts (success + fail) of the i'th mountain. We then do binary regression with the logit link to model the probability of success. This gives

1. Model for response: $Y_i \sim \text{Bin}(n_i, \pi_i)$,   for $i = 1, \ldots, 113$
2. Linear predictor: $\eta_i = \mathbf{x}_i^T \beta$
3. Link function: $\eta_i = \ln(\frac{\pi_i}{1 - \pi_i})$

where x_i is a p dimensional column vector of covariates for observation i, and β is the vector of regression parameters.

Part 2: Poisson regression - Eliteserien 2018

a)

We want to test if the assumption of independence between the goals made by the home and away teams is reasonable. To do this, we first load the data set and make a contingency table of all the results, with the goals of the home team on the rows, and goals of the away team on the columns. We get the following contingency table.

```
##      0  1  2 3 4+
## 0    8 18  3 1  1
## 1   19 26 15 5  3
## 2   10 14 13 4  1
## 3   13 10  7 2  0
## 4+   8  7  3 1  0
```

We then want to test if the number of goals for home and away team are independent. We do this by conducting *Pearson's χ^2 test* on the contingency table. The test poses the following hypotheses

H_0 : The sampling distributions are independently chi-squared distributed, H_1 : They are not independently chi-squared d

```
## 
##  Pearson's Chi-squared test
## 
## data:  contingency
## X-squared = 14.156, df = 16, p-value = 0.5871
```

We get a value of 14.146 for the test statistic, with a corresponding p-value of 0.5871. As this p-value is above any reasonable significance level, we keep the null hypothesis, and confirm that the goals scored by the home and away team are independent.

b)

We now make the current standings in the Eliteserie based on all the results in our data set, and get the following table.

```
##                    Team Played Won Drawn Lost For Against  GD Points
## 1         Rosenborg     24  16     4    4  43      20  23     52
## 2             Brann     24  14     6    4  36      23  13     48
## 3             Molde     24  13     4    7  48      30  18     43
## 4         Haugesund     24  12     5    7  36      28   8     41
## 5        Ranheim_TF     24  11     5    8  38      40  -2     38
## 6        Vaalerenga     24  10     6    8  35      37  -2     36
## 7               Odd     24   9     7    8  35      29   6     34
## 8           Tromsoe     24  10     3   11  35      33   2     33
## 9       Sarpsborg08     24   9     5   10  39      34   5     32
## 10      Kristiansund     24   8     7    9  32      35  -3     31
## 11       BodoeGlimt     24   6     9    9  28      30  -2     27
## 12      Stroemsgodset     24   6     8   10  38      38   0     26
## 13       Lillestroem     24   6     7   11  26      37 -11     25
## 14           Stabaek     24   5     8   11  29      43 -14     23
## 15             Start     24   6     5   13  24      42 -18     23
## 16 Sandefjord_Fotball     24   2     9   13  24      47 -23     15
```