PREMIER LEAGUE

Submitted to: Prof Amit S Ray

Harsh Mittal (22/42/HE/018) 4/29/24 Advanced Econometrics

Objective: I have tried to make use of Poisson regression techniques which is a useful modeling tool for count type dependent variables. Further, I have used Bayesian Poisson regression to increase the predictive power of the model. The dataset I have used is from the English Premier League database. Poisson distribution is a type of natural exponential family distribution, which uses GLM modeling for estimation, like the logistic model through the maximum likelihood method. Here I have presented the theoretical description, assumptions, and estimation technique of the model.

Natural Exponential Family

Suppose $y1, y2, \cdots, yn$ are independent observations

where yi has desnsity from natural exponential family

$f(yi|\theta i) = h(yi) \exp\{(\eta(\theta i)T(yi) - \psi(\theta i))\}$,

where $i = 1, 2, \cdots, n$.

- $\eta(\theta i)$ is known as canonical parameter
- $\psi(.)$ and $h(.)$ are known function

Poisson Distribution

Suppose $y_1, y_2, \cdots, y_n \sim Poisson(\theta_i)$

$$\begin{aligned} f(y_i|\theta_i) &= \frac{\theta_i^{y_i}}{y_i!}\exp\{-\theta_i\}, \\ &= \frac{1}{y_i!}\exp\{\log(\theta_i)y_i - \theta_i\} \end{aligned}$$

where $i = 1, 2, \cdots, n$.

- $h(yi) = 1/yi$
- $\eta(\theta i) = \log(\theta i)$
- $T(yi) = yi$
- $\psi(\theta i) = \theta i$

Hence, Poisson distribution is a special case of NEF.

Generalized Linear Model

1. Random Component $yi \sim NEF(\theta i)$ with pdf

$f(yi|\theta i) = h(yi) \exp\{(\eta(\theta i)T(yi) - \psi(\theta i))\}$, where i = 1, 2, · · · , n.

2. Link function: $\eta(\theta i) = zi$

3. Systematic component: $z_i = X_i^T \beta$

4. $\eta(\theta i) = X_i^T \beta$

Count (Poisson) Regression with GLM

1. Random Component $yi \sim \text{Poisson}(\theta i)$ with pdf

Suppose $y_1, y_2, \cdots, y_n \sim Poisson(\theta_i)$

$$\begin{aligned} f(y_i|\theta_i) &= \frac{\theta_i^{y_i}}{y_i!}\exp\{-\theta_i\}, \\ &= \frac{1}{y_i!}\exp\{\log(\theta_i)y_i - \theta_i\} \end{aligned}$$

where i = 1, 2, · · · , n.

2. Systematic component: $\eta(\theta i) = \log(\theta i) = X_i^T \beta$

Likelihood function of GLM

Negative log-Likelihood function of GLM

Negative log-Likelihood function of GLM

$$\begin{aligned} -\log L &= -\sum_{i=1}^{n} \log(f(y_i|\theta_i)) \\ &= -\sum_{i=1}^{n} \log(f(y_i|\eta^{-1}(\boldsymbol{x}_i^T\beta))) \end{aligned}$$

MLE of β of GLM

$$\hat{\beta}_{MLE} = \text{argmin}_{\beta}\left[-\sum_{i=1}^{n} \log(f(y_i|\eta^{-1}(\boldsymbol{x}_i^T\beta)))\right]$$

Key Assumptions in Poisson Regression:

1. Response Variable: The response variable consists of count data. For, e.g., the number of goals scored by the home team in our case.

2. Independence: The observations should be independent. Each count should be unrelated to the others in the dataset.

3. Constant rate of occurrence: The rate at which events occur is assumed to be constant over time and across different units of analysis. This means that the probability of an event occurring in a given time interval or unit of analysis is the same for all such intervals or units.

4. Linearity: The relationship between the expected counts of events (the mean) and the predictor variables is assumed to be linear on the log scale. This assumption is necessary because the Poisson regression model estimates the logarithm of the expected counts.

5. Mean-Variance Relationship: The variance of the counts should be equal to the mean. Overdispersion occurs when the variance is greater than the mean, indicating that there is more variability in the data than the Poisson distribution can account for. If overdispersion is present, alternative models such as the negative binomial regression may be more appropriate.

Data Description

Data has been taken from the English Premier League website for two time periods, 2020-21, and 2021-22. The first model without prior uses the later one, while the Bayesian model uses both datasets. Following is the data description, FTHG being the primary dependent variable:

Variable Definition

FTHG and HG = Full Time Home Team Goals

FTAG and AG = Full Time Away Team Goals

FTR and Res = Full Time Result (H=Home Win, D=Draw, A=Away Win)

HTHG = Half Time Home Team Goals

HTAG = Half Time Away Team Goals

HTR = Half Time Result (H=Home Win, D=Draw, A=Away Win)

Attendance = Crowd Attendance

Referee = Match Referee

HS = Home Team Shots

AS = Away Team Shots

HST = Home Team Shots on Target

AST = Away Team Shots on Target

HHW = Home Team Hit Woodwork

AHW = Away Team Hit Woodwork

HC = Home Team Corners

AC = Away Team Corners

HF = Home Team Fouls Committed

AF = Away Team Fouls Committed

HFKC = Home Team Free Kicks Conceded

AFKC = Away Team Free Kicks Conceded

HO = Home Team Offsides

AO = Away Team Offsides

HY = Home Team Yellow Cards

AY = Away Team Yellow Cards

HR = Home Team Red Cards

AR = Away Team Red Cards

HBP = Home Team Bookings Points (10 = yellow, 25 = red)

ABP = Away Team Bookings Points (10 = yellow, 25 = red)

B365H = Bet365 home win odds; B365D = Bet365 draw odds; B365A = Bet365 away win odds; BSH = Blue Square home win odds; BSD = Blue Square draw odds; BSA = Blue Square away win odds

Primary Model (without any prior)

$$\mathrm{FTHG} \sim \lambda$$
$$\log(\lambda)=Poisson(\lambda)\beta_0+\beta_1 HS+\beta_2 AS+\beta_3 HST+\beta_4 AST+\beta_5 HC+\beta_6 AC+\beta_7 B365H+\beta_8 B365A$$

English Premier League Data - Poisson Regression

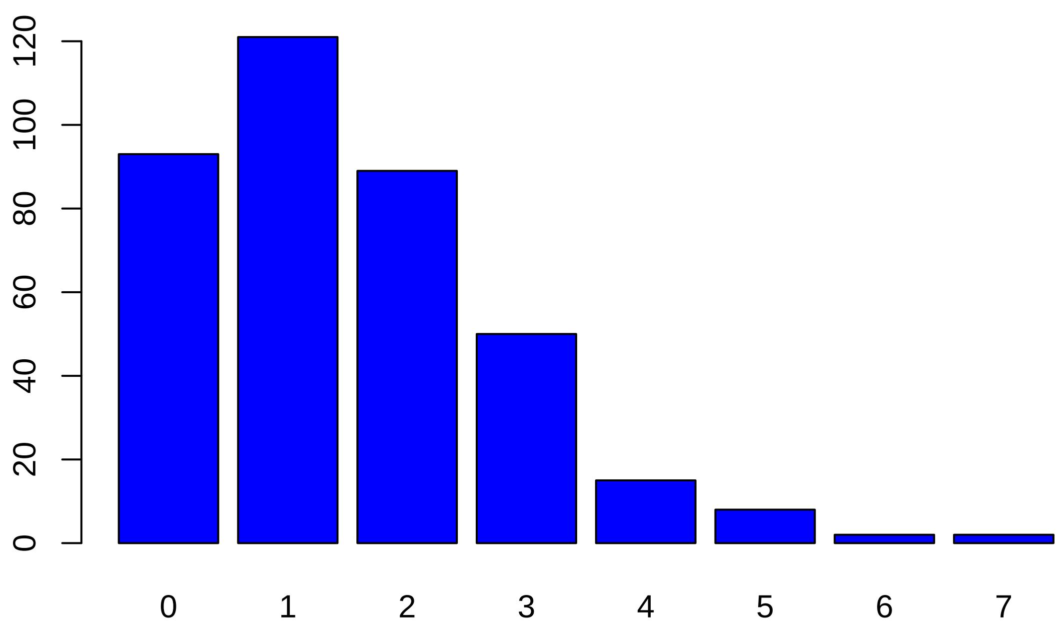
Harsh Mittal

Read the data

```r
data = read.csv("C:/Users/harsh.hm.mittal/Downloads/E0_21_22.csv")

table(data$FTHG)
```

```
##
##   0   1   2   3   4   5   6   7
##  93 121  89  50  15   8   2   2
```

```r
barplot(table(data$FTHG), col = 'Blue')
```



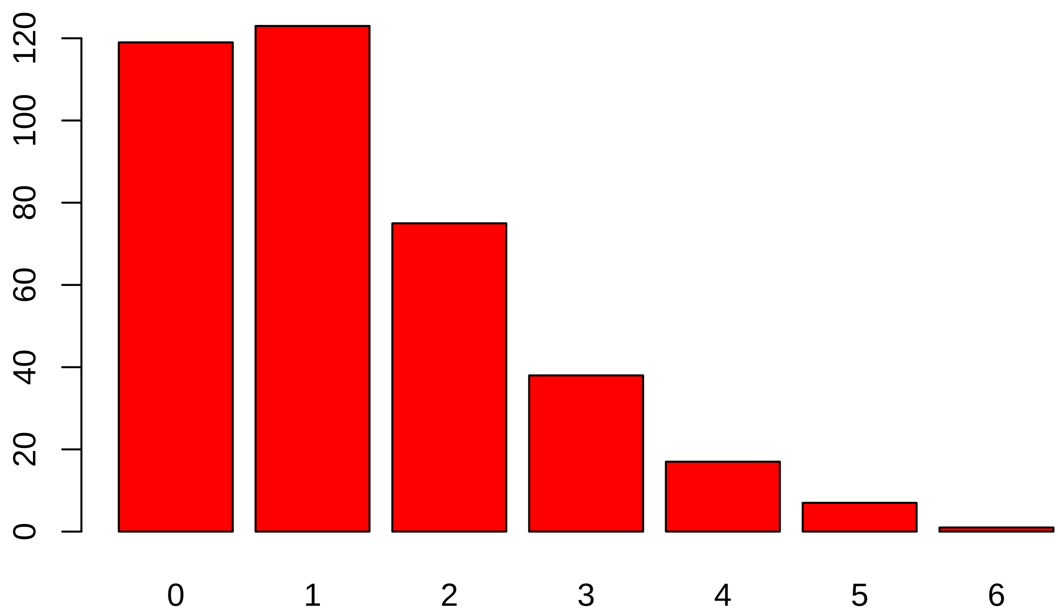

```r
mean(data$FTHG)
```

```
## [1] 1.513158
```

```r
table(data$FTAG)
```

```
##
##   0   1   2   3   4   5   6
## 119 123  75  38  17   7   1
```

```
barplot(table(data$FTAG), col = 'Red')
```




```
mean(data$FTAG)
```

```
## [1] 1.305263
```

##Modelling number of goals by Home Team as function of HS, HST, HC, HF, HY, HR ##FTHG ~ HS + HST+ HC + HF + HY+ HR

```
n = nrow(data)
m=ceiling(n*0.7)
set.seed(138)
train_id = sort(sample(1:n, m, replace = F))

df_train = data[train_id,]
df_test = data[-train_id,]
```

Poisson regression model

HS: Home Team Shots

HST: Home Team Shots on Target

HC: Home Team Corners

HF: Home Team Fouls Committed

HY: Home Team Yellow Cards

HR: Home Team Red Cards

```
mod1 = step(glm(FTHG ~ HS + HST+ HC + HF + HY+ HR
           ,data=df_train
           ,family = poisson(link = 'log')))
```

```
## Start:  AIC=748.45
## FTHG ~ HS + HST + HC + HF + HY + HR
##
##        Df Deviance    AIC
## - HF    1   230.81 746.46
## - HC    1   230.90 746.55
## - HS    1   230.96 746.60
## - HR    1   231.32 746.97
## <none>      230.80 748.45
## - HY    1   236.88 752.53
## - HST   1   291.61 807.26
##
## Step:  AIC=746.46
## FTHG ~ HS + HST + HC + HY + HR
##
##        Df Deviance    AIC
## - HC    1   230.91 744.56
## - HS    1   230.98 744.62
## - HR    1   231.33 744.98
## <none>      230.81 746.46
## - HY    1   237.33 750.98
## - HST   1   291.65 805.30
##
## Step:  AIC=744.56
## FTHG ~ HS + HST + HY + HR
##
##        Df Deviance    AIC
## - HS    1   231.29 742.94
## - HR    1   231.42 743.07
## <none>      230.91 744.56
## - HY    1   237.50 749.15
## - HST   1   293.62 805.26
##
## Step:  AIC=742.94
## FTHG ~ HST + HY + HR
##
##        Df Deviance    AIC
## - HR    1   231.74 741.38
```

```
## <none>      231.29 742.94
## - HY    1   237.52 747.17
## - HST   1   329.98 839.63
##
## Step:  AIC=741.38
## FTHG ~ HST + HY
##
##        Df Deviance    AIC
## <none>      231.74 741.38
## - HY    1   237.96 745.61
## - HST   1   333.90 841.55
```

```r
summary(mod1)
```

```
##
## Call:
## glm(formula = FTHG ~ HST + HY, family = poisson(link = "log"),
##     data = df_train)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.11917  -1.09148  -0.09127   0.57635   2.47377
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.29735    0.13929  -2.135   0.0328 *
## HST          0.16587    0.01566  10.591   <2e-16 ***
## HY          -0.11036    0.04491  -2.457   0.0140 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 364.48  on 265  degrees of freedom
## Residual deviance: 231.74  on 263  degrees of freedom
## AIC: 741.38
##
## Number of Fisher Scoring iterations: 5
```

```r
mod2 = step(glm(FTHG ~ HS + HST+ HC + HF + HY+ HR
                + AS + AST + AC + AF + AY+ AR
                + I(HST^2) + I (HST^3) + I(HST^4) + I(HST^5)
                + I(AST^2) + I (AST^3) + I(AST^4) + I(AST^5)
                ,data=df_train
                ,family = poisson(link = 'log')))
```

```
## Start:  AIC=742.51
## FTHG ~ HS + HST + HC + HF + HY + HR + AS + AST + AC + AF + AY +
##     AR + I(HST^2) + I(HST^3) + I(HST^4) + I(HST^5) + I(AST^2) +
##     I(AST^3) + I(AST^4) + I(AST^5)
##
##            Df Deviance    AIC
## - I(HST^4)  1   196.86 740.51
## - I(HST^5)  1   196.87 740.52
## - I(HST^3)  1   196.88 740.53
```

```
## - AS        1   196.91 740.56
## - AC        1   196.93 740.58
## - AF        1   196.95 740.59
## - I(AST^4)  1   196.98 740.63
## - I(AST^5)  1   196.98 740.63
## - I(AST^3)  1   197.00 740.65
## - I(AST^2)  1   197.07 740.72
## - AR        1   197.11 740.76
## - I(HST^2)  1   197.11 740.76
## - HR        1   197.28 740.93
## - AST       1   197.32 740.97
## - HF        1   197.32 740.97
## - HC        1   197.41 741.06
## - HS        1   197.50 741.14
## - AY        1   197.70 741.35
## <none>          196.86 742.51
## - HST       1   199.34 742.99
## - HY        1   202.26 745.91
##
## Step:  AIC=740.51
## FTHG ~ HS + HST + HC + HF + HY + HR + AS + AST + AC + AF + AY +
##     AR + I(HST^2) + I(HST^3) + I(HST^5) + I(AST^2) + I(AST^3) +
##     I(AST^4) + I(AST^5)
##
##            Df Deviance    AIC
## - AS        1   196.91 738.56
## - AC        1   196.93 738.58
## - AF        1   196.95 738.60
## - I(AST^4)  1   196.98 738.63
## - I(AST^5)  1   196.98 738.63
## - I(AST^3)  1   197.00 738.65
## - I(AST^2)  1   197.07 738.72
## - I(HST^5)  1   197.08 738.73
## - AR        1   197.11 738.76
## - HR        1   197.28 738.93
## - AST       1   197.32 738.97
## - HF        1   197.32 738.97
## - HC        1   197.42 739.07
## - HS        1   197.50 739.15
## - AY        1   197.71 739.35
## - I(HST^3)  1   198.27 739.92
## <none>          196.86 740.51
## - I(HST^2)  1   200.46 742.11
## - HY        1   202.31 743.96
## - HST       1   208.83 750.48
##
## Step:  AIC=738.56
## FTHG ~ HS + HST + HC + HF + HY + HR + AST + AC + AF + AY + AR +
##     I(HST^2) + I(HST^3) + I(HST^5) + I(AST^2) + I(AST^3) + I(AST^4) +
##     I(AST^5)
##
##            Df Deviance    AIC
## - AC        1   196.95 736.60
## - AF        1   196.99 736.63
```

```
## - I(AST^4)  1   197.02 736.67
## - I(AST^5)  1   197.02 736.67
## - I(AST^3)  1   197.04 736.69
## - I(HST^5)  1   197.10 736.75
## - I(AST^2)  1   197.10 736.75
## - AR        1   197.17 736.82
## - HR        1   197.35 737.00
## - AST       1   197.36 737.01
## - HF        1   197.37 737.01
## - HC        1   197.43 737.08
## - HS        1   197.52 737.17
## - AY        1   197.75 737.40
## - I(HST^3)  1   198.27 737.92
## <none>          196.91 738.56
## - I(HST^2)  1   200.46 740.11
## - HY        1   202.52 742.17
## - HST       1   208.83 748.48
##
## Step:  AIC=736.6
## FTHG ~ HS + HST + HC + HF + HY + HR + AST + AF + AY + AR + I(HST^2) +
##     I(HST^3) + I(HST^5) + I(AST^2) + I(AST^3) + I(AST^4) + I(AST^5)
##
##            Df Deviance    AIC
## - AF        1   197.03 734.67
## - I(AST^4)  1   197.06 734.71
## - I(AST^5)  1   197.06 734.71
## - I(AST^3)  1   197.08 734.73
## - I(AST^2)  1   197.14 734.79
## - I(HST^5)  1   197.14 734.79
## - AR        1   197.19 734.84
## - AST       1   197.39 735.04
## - HR        1   197.42 735.07
## - HF        1   197.43 735.07
## - HC        1   197.50 735.15
## - HS        1   197.61 735.25
## - AY        1   197.76 735.41
## - I(HST^3)  1   198.31 735.96
## <none>          196.95 736.60
## - I(HST^2)  1   200.49 738.14
## - HY        1   202.63 740.27
## - HST       1   208.83 746.48
##
## Step:  AIC=734.67
## FTHG ~ HS + HST + HC + HF + HY + HR + AST + AY + AR + I(HST^2) +
##     I(HST^3) + I(HST^5) + I(AST^2) + I(AST^3) + I(AST^4) + I(AST^5)
##
##            Df Deviance    AIC
## - I(AST^4)  1   197.12 732.77
## - I(AST^5)  1   197.12 732.77
## - I(AST^3)  1   197.14 732.78
## - I(AST^2)  1   197.20 732.85
## - I(HST^5)  1   197.21 732.85
## - AR        1   197.22 732.87
## - AST       1   197.46 733.11
```

```
## - HR        1   197.48 733.13
## - HC        1   197.54 733.19
## - HF        1   197.57 733.22
## - HS        1   197.67 733.32
## - AY        1   197.76 733.41
## - I(HST^3)  1   198.36 734.00
## <none>          197.03 734.67
## - I(HST^2)  1   200.53 736.18
## - HY        1   202.63 738.28
## - HST       1   208.86 744.51
##
## Step:  AIC=732.77
## FTHG ~ HS + HST + HC + HF + HY + HR + AST + AY + AR + I(HST^2) +
##     I(HST^3) + I(HST^5) + I(AST^2) + I(AST^3) + I(AST^5)
##
##            Df Deviance    AIC
## - I(AST^5)  1   197.12 730.77
## - I(AST^3)  1   197.16 730.81
## - I(AST^2)  1   197.28 730.93
## - I(HST^5)  1   197.29 730.94
## - AR        1   197.32 730.97
## - AST       1   197.62 731.26
## - HR        1   197.64 731.29
## - HF        1   197.66 731.31
## - HS        1   197.71 731.36
## - HC        1   197.73 731.38
## - AY        1   197.85 731.50
## - I(HST^3)  1   198.42 732.07
## <none>          197.12 732.77
## - I(HST^2)  1   200.57 734.21
## - HY        1   202.63 736.28
## - HST       1   208.86 742.51
##
## Step:  AIC=730.77
## FTHG ~ HS + HST + HC + HF + HY + HR + AST + AY + AR + I(HST^2) +
##     I(HST^3) + I(HST^5) + I(AST^2) + I(AST^3)
##
##            Df Deviance    AIC
## - I(HST^5)  1   197.29 728.94
## - AR        1   197.33 728.97
## - HR        1   197.64 729.29
## - HF        1   197.67 729.32
## - HS        1   197.71 729.36
## - HC        1   197.74 729.39
## - AY        1   197.85 729.50
## - I(AST^3)  1   198.01 729.65
## - I(AST^2)  1   198.13 729.78
## - I(HST^3)  1   198.42 730.07
## - AST       1   198.47 730.12
## <none>          197.12 730.77
## - I(HST^2)  1   200.57 732.22
## - HY        1   202.63 734.28
## - HST       1   208.88 740.53
##
```

```
## Step:  AIC=728.94
## FTHG ~ HS + HST + HC + HF + HY + HR + AST + AY + AR + I(HST^2) +
##     I(HST^3) + I(AST^2) + I(AST^3)
##
##            Df Deviance    AIC
## - AR        1   197.54 727.19
## - HR        1   197.81 727.45
## - HF        1   197.84 727.49
## - HS        1   197.88 727.52
## - HC        1   197.93 727.58
## - AY        1   198.06 727.71
## - I(AST^3)  1   198.35 727.99
## - I(AST^2)  1   198.53 728.18
## - AST       1   198.95 728.60
## <none>          197.29 728.94
## - HY        1   203.03 732.68
## - I(HST^3)  1   211.11 740.76
## - I(HST^2)  1   216.35 746.00
## - HST       1   234.10 763.75
##
## Step:  AIC=727.19
## FTHG ~ HS + HST + HC + HF + HY + HR + AST + AY + I(HST^2) + I(HST^3) +
##     I(AST^2) + I(AST^3)
##
##            Df Deviance    AIC
## - HF        1   198.00 725.65
## - HS        1   198.07 725.72
## - HR        1   198.07 725.72
## - HC        1   198.19 725.84
## - AY        1   198.34 725.99
## - I(AST^3)  1   198.65 726.30
## - I(AST^2)  1   198.85 726.50
## - AST       1   199.34 726.98
## <none>          197.54 727.19
## - HY        1   203.41 731.06
## - I(HST^3)  1   211.17 738.82
## - I(HST^2)  1   216.39 744.04
## - HST       1   234.12 761.76
##
## Step:  AIC=725.65
## FTHG ~ HS + HST + HC + HY + HR + AST + AY + I(HST^2) + I(HST^3) +
##     I(AST^2) + I(AST^3)
##
##            Df Deviance    AIC
## - HS        1   198.45 724.10
## - HR        1   198.58 724.23
## - HC        1   198.67 724.32
## - AY        1   198.69 724.34
## - I(AST^3)  1   199.08 724.73
## - I(AST^2)  1   199.27 724.92
## - AST       1   199.70 725.35
## <none>          198.00 725.65
## - HY        1   205.42 731.07
## - I(HST^3)  1   211.75 737.40
```

```
## - I(HST^2)  1   216.82 742.47
## - HST       1   234.36 760.01
##
## Step:  AIC=724.1
## FTHG ~ HST + HC + HY + HR + AST + AY + I(HST^2) + I(HST^3) +
##     I(AST^2) + I(AST^3)
##
##            Df Deviance    AIC
## - HR        1   198.98 722.63
## - AY        1   199.09 722.74
## - I(AST^3)  1   199.57 723.22
## - I(AST^2)  1   199.76 723.41
## - HC        1   199.90 723.55
## - AST       1   200.11 723.76
## <none>          198.45 724.10
## - HY        1   205.45 729.09
## - I(HST^3)  1   212.03 735.68
## - I(HST^2)  1   217.08 740.73
## - HST       1   234.38 758.03
##
## Step:  AIC=722.63
## FTHG ~ HST + HC + HY + AST + AY + I(HST^2) + I(HST^3) + I(AST^2) +
##     I(AST^3)
##
##            Df Deviance    AIC
## - AY        1   199.53 721.17
## - I(AST^3)  1   199.92 721.57
## - I(AST^2)  1   200.09 721.74
## - HC        1   200.38 722.03
## - AST       1   200.44 722.09
## <none>          198.98 722.63
## - HY        1   205.91 727.56
## - I(HST^3)  1   212.38 734.03
## - I(HST^2)  1   217.44 739.09
## - HST       1   234.89 756.54
##
## Step:  AIC=721.17
## FTHG ~ HST + HC + HY + AST + I(HST^2) + I(HST^3) + I(AST^2) +
##     I(AST^3)
##
##            Df Deviance    AIC
## - I(AST^3)  1   200.48 720.13
## - I(AST^2)  1   200.64 720.29
## - HC        1   200.93 720.58
## - AST       1   200.95 720.60
## <none>          199.53 721.17
## - HY        1   205.99 725.64
## - I(HST^3)  1   213.43 733.08
## - I(HST^2)  1   218.48 738.13
## - HST       1   235.99 755.64
##
## Step:  AIC=720.13
## FTHG ~ HST + HC + HY + AST + I(HST^2) + I(HST^3) + I(AST^2)
##
```

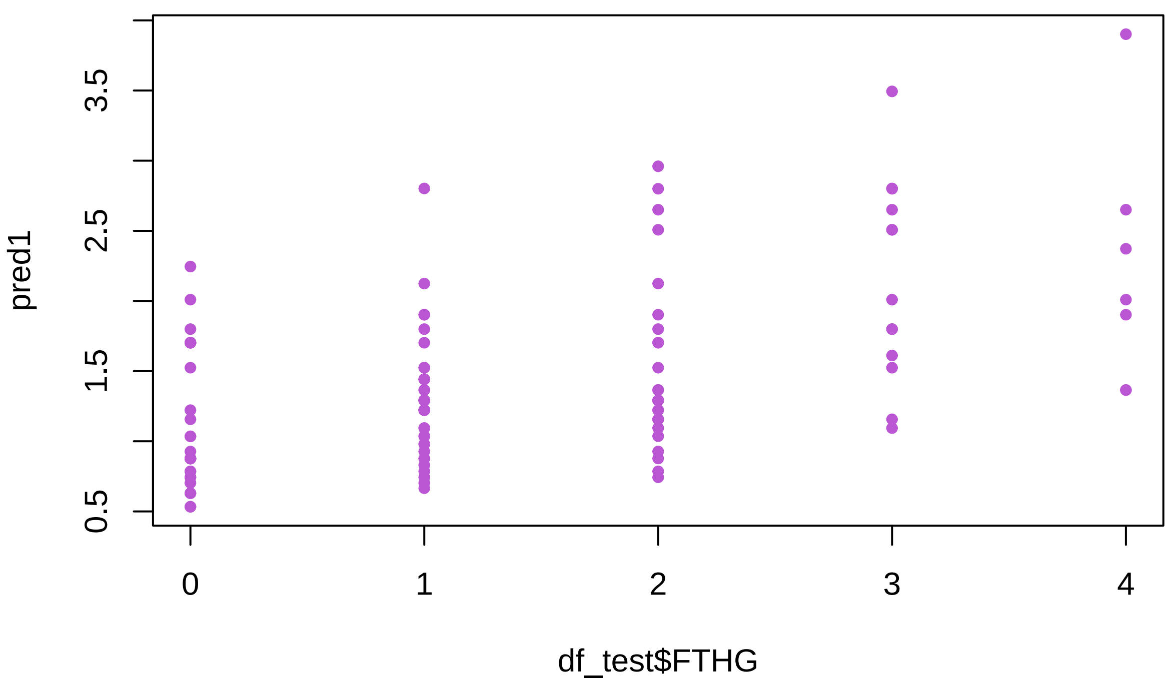
```
##             Df Deviance    AIC
## - I(AST^2)  1   200.70 718.35
## - AST       1   200.98 718.63
## - HC        1   201.93 719.58
## <none>          200.48 720.13
## - HY        1   207.62 725.26
## - I(HST^3)  1   213.76 731.41
## - I(HST^2)  1   218.73 736.38
## - HST       1   236.10 753.75
##
## Step:  AIC=718.35
## FTHG ~ HST + HC + HY + AST + I(HST^2) + I(HST^3)
##
##             Df Deviance    AIC
## - AST       1   201.19 716.84
## - HC        1   202.09 717.74
## <none>          200.70 718.35
## - HY        1   207.84 723.48
## - I(HST^3)  1   213.92 729.56
## - I(HST^2)  1   218.86 734.51
## - HST       1   236.23 751.88
##
## Step:  AIC=716.84
## FTHG ~ HST + HC + HY + I(HST^2) + I(HST^3)
##
##             Df Deviance    AIC
## - HC        1   202.41 716.06
## <none>          201.19 716.84
## - HY        1   208.58 722.23
## - I(HST^3)  1   214.20 727.85
## - I(HST^2)  1   219.13 732.78
## - HST       1   236.67 750.32
##
## Step:  AIC=716.06
## FTHG ~ HST + HY + I(HST^2) + I(HST^3)
##
##             Df Deviance    AIC
## <none>          202.41 716.06
## - HY        1   209.32 720.97
## - I(HST^3)  1   216.52 728.16
## - I(HST^2)  1   221.16 732.81
## - HST       1   237.89 749.54
```

```r
summary(mod2)
```

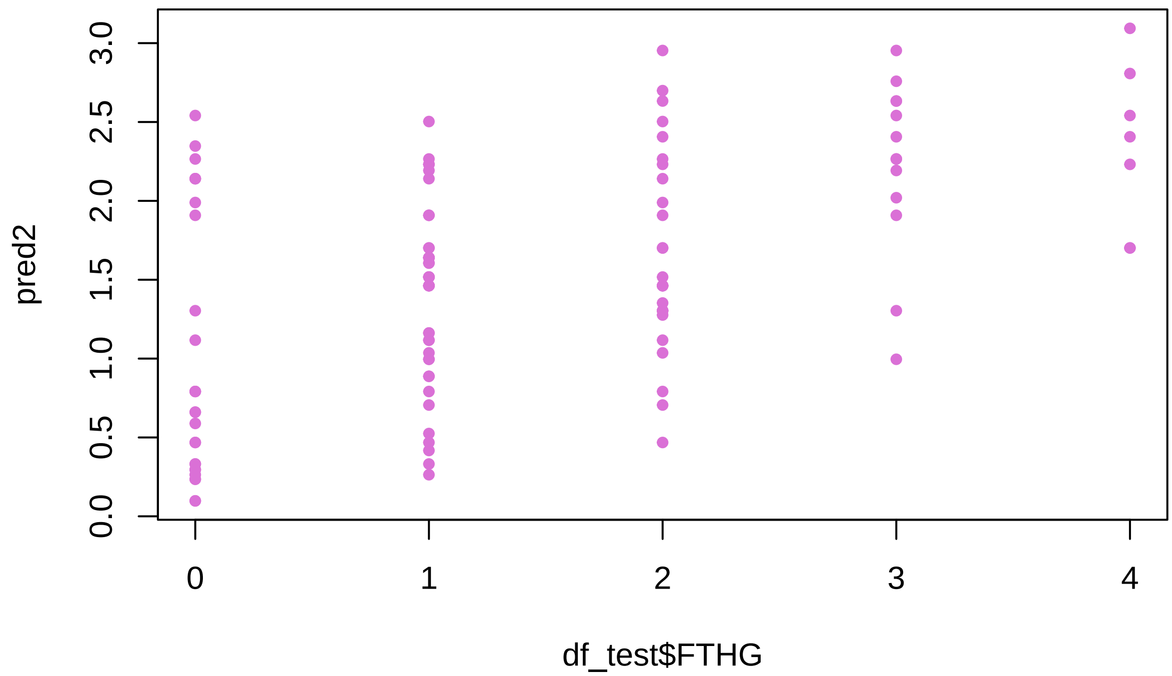
```
##
## Call:
## glm(formula = FTHG ~ HST + HY + I(HST^2) + I(HST^3), family = poisson(link = "log"),
##     data = df_train)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.2374  -0.7603  -0.1690   0.4645   2.3331
##
## Coefficients:
```

```
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.979650   0.401728  -4.928 8.31e-07 ***
## HST          0.976904   0.183659   5.319 1.04e-07 ***
## HY          -0.114787   0.044369  -2.587 0.009679 **
## I(HST^2)    -0.104901   0.025698  -4.082 4.46e-05 ***
## I(HST^3)     0.003830   0.001053   3.637 0.000276 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 364.48  on 265  degrees of freedom
## Residual deviance: 202.41  on 261  degrees of freedom
## AIC: 716.06
##
## Number of Fisher Scoring iterations: 5
```

```r
pred1 = predict(mod1, newdata = df_test, type = "response")

plot(df_test$FTHG, pred1, pch = 20, col = 'MediumOrchid')
```




```r
R2_approx = cor(df_test$FTHG, pred1)^2


pred2 = predict(mod2, newdata = df_test, type = "response")

plot(df_test$FTHG, pred2, pch = 20, col = 'Orchid')
```

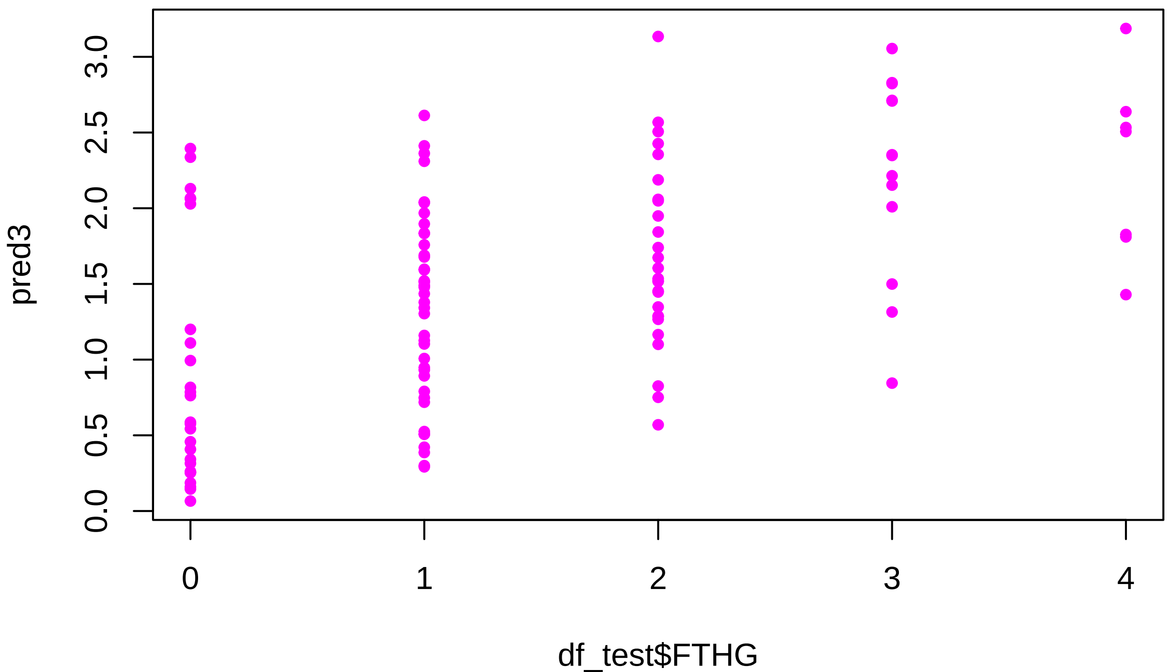



```r
R2_approx = cor(df_test$FTHG, pred2)^2

mod3 = update(mod2, .~. + B365H)
summary(mod3)
```

```
##
## Call:
## glm(formula = FTHG ~ HST + HY + I(HST^2) + I(HST^3) + B365H,
##     family = poisson(link = "log"), data = df_train)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.1149  -0.7587  -0.0925   0.4686   2.2460
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.453960   0.442355  -3.287 0.001013 **
## HST          0.874387   0.184870   4.730 2.25e-06 ***
## HY          -0.108373   0.044539  -2.433 0.014966 *
## I(HST^2)    -0.094838   0.025660  -3.696 0.000219 ***
## I(HST^3)     0.003491   0.001050   3.324 0.000886 ***
## B365H       -0.098973   0.040623  -2.436 0.014834 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
```

```
##
##     Null deviance: 364.48  on 265  degrees of freedom
## Residual deviance: 195.49  on 260  degrees of freedom
## AIC: 711.14
##
## Number of Fisher Scoring iterations: 5
```

```r
pred3 = predict(mod3, newdata = df_test, type = "response")
plot(df_test$FTHG, pred3, pch = 20, col = 'Magenta')
```




```r
R2_approx = cor(df_test$FTHG, pred3)^2

c(cor(df_test$FTHG, pred1)^2
  ,cor(df_test$FTHG, pred2)^2
  ,cor(df_test$FTHG, pred3)^2)
```

```
## [1] 0.2752409 0.2900639 0.3207744
```

Assumptions Check

Independence

```r
library(randtests)
```

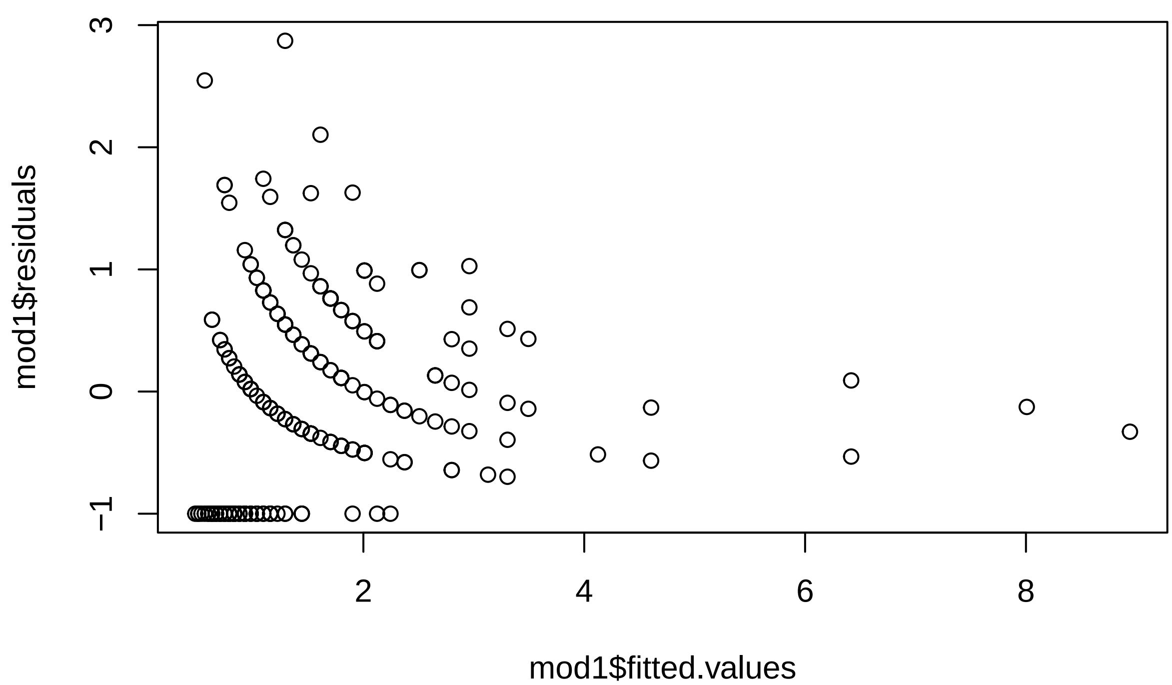
```
## Warning: package 'randtests' was built under R version 4.2.3
```

```r
bartels.rank.test(mod1$residuals)
```

```
##
##  Bartels Ratio Test
##
## data:  mod1$residuals
## statistic = -0.67192, n = 266, p-value = 0.5016
## alternative hypothesis: nonrandomness
```

#Linearity

```r
plot(mod1$fitted.values, mod1$residuals)
```




```r
cor(mod1$fitted.values, mod1$residuals)
```

```
## [1] 0.07376815
```

Since, these models have weak predictive power, hence we'll now proceed with Bayesian Poisson Regression Models in which we use prior.

HARSH MITTAL

Likelihood and Bayesian Poisson Regression using Julia & CRRao

- Maximum Likelihood Methods
- Ridge Prior
- Laplace Prior
- Cauchy Prior
- T-Distributed Prior
- Horse Shoe Prior

In [1]:
```
using CSV, DataFrames, CRRao
using Statistics;
```

In [2]:
```
df_train = DataFrame(CSV.File("E0_20_21.csv"));
```

Model

$$FTHG \quad Poisson(\lambda) \tag{1}$$
$$\log(\lambda) = \beta_0 + \beta_1 HS + \beta_2 AS + \beta_3 HST + \beta_4 AST + \beta_5 HC + \beta_6 AC + \beta_7 B365H + \beta_8 B365A \tag{2}$$

MLE Estimates

In [3]:
```
using CRRao, StatsModels

mod1 = fit(@formula(FTHG ~ HS + AS + HST + AST + HC + AC +B365H + B365A),df_train,PoissonRegression())
```

Out[3]:
```
Model Class: Poisson Regression
Likelihood Mode: Poison
Link Function: Log
Computing Method: Optimization
──────────────────────────────────────────────────────────────────────────────
                  Coef.  Std. Error      z  Pr(>|z|)   Lower 95%    Upper 95%
──────────────────────────────────────────────────────────────────────────────
(Intercept)  -0.352743    0.236089   -1.49    0.1351  -0.815468     0.109983
HS           -0.034822    0.0141776  -2.46    0.0140  -0.0626096   -0.00703447
AS           -0.023019    0.0139061  -1.66    0.0979  -0.0502744    0.00423639
HST           0.238636    0.0229936  10.38    <1e-24   0.193569     0.283703
AST           0.0255725   0.0271946   0.94    0.3470  -0.0277279    0.0788729
HC           -0.0438155   0.0186827  -2.35    0.0190  -0.0804331   -0.00719804
AC            0.0344269   0.0193173   1.78    0.0747  -0.00343431   0.0722882
B365H        -0.0248462   0.0300651  -0.83    0.4086  -0.0837727    0.0340803
B365A         0.0368102   0.0145484   2.53    0.0114   0.00829595   0.0653245
──────────────────────────────────────────────────────────────────────────────
```

Bayesian Poisson Regression with Ridge Prior

In [4]:
```
mod_ridge = fit(@formula(FTHG ~ HS + AS + HST + AST + HC + AC +B365H + B365A)
                   ,df_train,PoissonRegression()
                   ,Prior_Ridge())
```

```
┌ Info: Found initial step size
└   ϵ = 0.00078125
Sampling: 100%|█████████████████████████████████████████| Time: 0:00:02
```

Out[4]:

```
Formula: FTHG ~ 1 + HS + AS + HST + AST + HC + AC + B365H + B365A
Link: CRRao.Identity(CRRao.Identity_Link)
Chain: Chains MCMC chain (1000×22×1 Array{Float64, 3}):

Iterations        = 501:1:1500
Number of chains  = 1
Samples per chain = 1000
Wall duration     = 9.67 seconds
Compute duration  = 9.67 seconds
parameters        = λ, β[1], β[2], β[3], β[4], β[5], β[6], β[7], β[8], β[9]
internals         = lp, n_steps, is_accept, acceptance_rate, log_density, hamiltonian_energy, hamiltonian_ener
gy_error, max_hamiltonian_energy_error, tree_depth, numerical_error, step_size, nom_step_size

Summary Statistics
  parameters      mean       std   naive_se      mcse         ess      rhat   ⋯
      Symbol   Float64   Float64    Float64   Float64     Float64   Float64   ⋯

           λ    0.1054    0.0339     0.0011    0.0014    464.9617    1.0000   ⋯
        β[1]   -0.0623    0.1052     0.0033    0.0047    397.6786    1.0076   ⋯
        β[2]   -0.0346    0.0131     0.0004    0.0004    747.5306    0.9995   ⋯
        β[3]   -0.0267    0.0127     0.0004    0.0005    714.2560    0.9998   ⋯
        β[4]    0.2259    0.0226     0.0007    0.0006    829.1529    0.9990   ⋯
        β[5]    0.0169    0.0261     0.0008    0.0009    752.0529    0.9990   ⋯
        β[6]   -0.0481    0.0184     0.0006    0.0007    911.7310    0.9992   ⋯
        β[7]    0.0282    0.0188     0.0006    0.0005   1093.4051    1.0027   ⋯
        β[8]   -0.0400    0.0265     0.0008    0.0009    868.9328    0.9992   ⋯
        β[9]    0.0274    0.0127     0.0004    0.0004   1028.8591    1.0054   ⋯
                                                           1 column omitted

Quantiles
  parameters      2.5%     25.0%     50.0%     75.0%     97.5%
      Symbol   Float64   Float64   Float64   Float64   Float64

           λ    0.0605    0.0826    0.1000    0.1201    0.1981
        β[1]   -0.2970   -0.1122   -0.0517    0.0018    0.1119
        β[2]   -0.0605   -0.0437   -0.0346   -0.0253   -0.0088
        β[3]   -0.0516   -0.0353   -0.0262   -0.0183   -0.0018
        β[4]    0.1797    0.2110    0.2257    0.2414    0.2701
        β[5]   -0.0338   -0.0005    0.0165    0.0334    0.0710
        β[6]   -0.0822   -0.0608   -0.0483   -0.0357   -0.0120
        β[7]   -0.0111    0.0152    0.0283    0.0416    0.0646
        β[8]   -0.0914   -0.0575   -0.0395   -0.0219    0.0113
        β[9]    0.0010    0.0196    0.0277    0.0363    0.0508
```

Bayesian Poisson Regression with Laplace Prior

In [5]:

```
mod_laplace = fit(@formula(FTHG ~ HS + AS + HST + AST + HC + AC +B365H + B365A)
                  ,df_train,PoissonRegression()
                  ,Prior_Laplace())
```

```
┌ Info: Found initial step size
└   ϵ = 0.00625
Sampling: 100%|█████████████████████████████████████████| Time: 0:00:02
```

Out[5]:
```
Formula: FTHG ~ 1 + HS + AS + HST + AST + HC + AC + B365H + B365A
Link: CRRao.Identity(CRRao.Identity_Link)
Chain: Chains MCMC chain (1000×22×1 Array{Float64, 3}):

Iterations        = 501:1:1500
Number of chains  = 1
Samples per chain = 1000
Wall duration     = 4.83 seconds
Compute duration  = 4.83 seconds
parameters        = λ, β[1], β[2], β[3], β[4], β[5], β[6], β[7], β[8], β[9]
internals         = lp, n_steps, is_accept, acceptance_rate, log_density, hamiltonian_energy, hamiltonian_ener
gy_error, max_hamiltonian_energy_error, tree_depth, numerical_error, step_size, nom_step_size

Summary Statistics
  parameters      mean       std   naive_se      mcse        ess      rhat   e ⋯
      Symbol   Float64   Float64    Float64   Float64    Float64   Float64     ⋯

           λ    0.0787    0.0344     0.0011    0.0015   561.0030    0.9995     ⋯
        β[1]   -0.0708    0.1243     0.0039    0.0071   309.8460    1.0003     ⋯
        β[2]   -0.0360    0.0137     0.0004    0.0005   675.3930    0.9995     ⋯
        β[3]   -0.0244    0.0128     0.0004    0.0005   674.2893    1.0009     ⋯
        β[4]    0.2302    0.0242     0.0008    0.0010   722.7525    1.0015     ⋯
        β[5]    0.0114    0.0243     0.0008    0.0009   591.8113    0.9992     ⋯
        β[6]   -0.0462    0.0191     0.0006    0.0005   873.4254    0.9996     ⋯
        β[7]    0.0231    0.0181     0.0006    0.0006   831.4847    0.9999     ⋯
        β[8]   -0.0332    0.0270     0.0009    0.0010   811.8887    0.9992     ⋯
        β[9]    0.0264    0.0137     0.0004    0.0004   638.7690    1.0011     ⋯
                                                             1 column omitted

Quantiles
  parameters      2.5%     25.0%     50.0%     75.0%     97.5%
      Symbol   Float64   Float64   Float64   Float64   Float64

           λ    0.0375    0.0563    0.0716    0.0912    0.1619
        β[1]   -0.4111   -0.1190   -0.0375    0.0066    0.0998
        β[2]   -0.0628   -0.0452   -0.0361   -0.0260   -0.0097
        β[3]   -0.0519   -0.0324   -0.0242   -0.0152   -0.0006
        β[4]    0.1848    0.2127    0.2300    0.2466    0.2777
        β[5]   -0.0353   -0.0049    0.0102    0.0279    0.0623
        β[6]   -0.0838   -0.0595   -0.0461   -0.0337   -0.0062
        β[7]   -0.0095    0.0100    0.0231    0.0355    0.0594
        β[8]   -0.0859   -0.0526   -0.0321   -0.0154    0.0201
        β[9]    0.0004    0.0165    0.0262    0.0360    0.0533
```

Bayesian Poisson Regression with Cauchy Prior

In [6]:
```
mod_Cauchy = fit(@formula(FTHG ~ HS + AS + HST + AST + HC + AC +B365H + B365A)
                   ,df_train,PoissonRegression()
                   ,Prior_Cauchy())
```

```
┌ Info: Found initial step size
└   ϵ = 7.450580596923829e-10
Sampling: 100%|██████████████████████████████████████████| Time: 0:00:02
```

Out[6]:

```
Formula: FTHG ~ 1 + HS + AS + HST + AST + HC + AC + B365H + B365A
Link: CRRao.Identity(CRRao.Identity_Link)
Chain: Chains MCMC chain (1000×22×1 Array{Float64, 3}):

Iterations        = 501:1:1500
Number of chains  = 1
Samples per chain = 1000
Wall duration     = 5.11 seconds
Compute duration  = 5.11 seconds
parameters        = λ, β[1], β[2], β[3], β[4], β[5], β[6], β[7], β[8], β[9]
internals         = lp, n_steps, is_accept, acceptance_rate, log_density, hamiltonian_energy, hamiltonian_ener
gy_error, max_hamiltonian_energy_error, tree_depth, numerical_error, step_size, nom_step_size

Summary Statistics
  parameters      mean       std   naive_se      mcse         ess      rhat   ⋯
      Symbol   Float64   Float64    Float64   Float64     Float64   Float64   ⋯

           λ    0.1627    0.0590     0.0019    0.0020    896.0429    0.9995   ⋯
        β[1]   -0.1611    0.1845     0.0058    0.0061    864.8895    1.0003   ⋯
        β[2]   -0.0377    0.0136     0.0004    0.0004    889.7947    0.9995   ⋯
        β[3]   -0.0258    0.0138     0.0004    0.0004    972.0278    0.9999   ⋯
        β[4]    0.2358    0.0220     0.0007    0.0007    986.9997    0.9990   ⋯
        β[5]    0.0212    0.0270     0.0009    0.0005   1158.4883    1.0004   ⋯
        β[6]   -0.0461    0.0187     0.0006    0.0005   1135.6974    1.0003   ⋯
        β[7]    0.0307    0.0194     0.0006    0.0004   1223.4360    0.9999   ⋯
        β[8]   -0.0356    0.0285     0.0009    0.0009   1156.0803    0.9997   ⋯
        β[9]    0.0307    0.0134     0.0004    0.0004   1101.4153    1.0004   ⋯
                                                          1 column omitted

Quantiles
  parameters      2.5%     25.0%     50.0%     75.0%     97.5%
      Symbol   Float64   Float64   Float64   Float64   Float64

           λ    0.0841    0.1227    0.1507    0.1898    0.3116
        β[1]   -0.5707   -0.2686   -0.1351   -0.0272    0.1458
        β[2]   -0.0638   -0.0469   -0.0382   -0.0281   -0.0118
        β[3]   -0.0527   -0.0347   -0.0258   -0.0169    0.0019
        β[4]    0.1914    0.2220    0.2361    0.2506    0.2784
        β[5]   -0.0290    0.0015    0.0217    0.0401    0.0743
        β[6]   -0.0836   -0.0594   -0.0459   -0.0328   -0.0094
        β[7]   -0.0084    0.0180    0.0309    0.0436    0.0674
        β[8]   -0.0898   -0.0551   -0.0355   -0.0164    0.0185
        β[9]    0.0049    0.0214    0.0312    0.0402    0.0549
```

Bayesian Poisson Regression with TDistributed Prior

In [7]:

```
mod_TDist = fit(@formula(FTHG ~ HS + AS + HST + AST + HC + AC +B365H + B365A)
                    ,df_train,PoissonRegression()
                    ,Prior_TDist(),0.95,10000)
```

```
┌ Info: Found initial step size
└   ϵ = 0.0125
Sampling: 100%|█████████████████████████████████████████| Time: 0:00:14
```

Out[7]:
```
Formula: FTHG ~ 1 + HS + AS + HST + AST + HC + AC + B365H + B365A
Link: CRRao.Identity(CRRao.Identity_Link)
Chain: Chains MCMC chain (10000×23×1 Array{Float64, 3}):

Iterations        = 1001:1:11000
Number of chains  = 1
Samples per chain = 10000
Wall duration     = 19.01 seconds
Compute duration  = 19.01 seconds
parameters        = λ, ν, β[1], β[2], β[3], β[4], β[5], β[6], β[7], β[8], β[9]
internals         = lp, n_steps, is_accept, acceptance_rate, log_density, hamiltonian_energy, hamiltonian_ener
gy_error, max_hamiltonian_energy_error, tree_depth, numerical_error, step_size, nom_step_size

Summary Statistics
  parameters      mean        std   naive_se      mcse         ess      rhat  ⋯
      Symbol   Float64    Float64    Float64   Float64     Float64   Float64  ⋯

           λ    0.1657     0.0583     0.0006    0.0006   6946.7029    1.0001  ⋯
           ν   37.4676   816.4755     8.1648    9.8767   6717.1613    0.9999  ⋯
        β[1]   -0.1359     0.1649     0.0016    0.0020   6949.4095    1.0003  ⋯
        β[2]   -0.0371     0.0135     0.0001    0.0001   8245.6082    0.9999  ⋯
        β[3]   -0.0265     0.0134     0.0001    0.0002   8096.6953    1.0002  ⋯
        β[4]    0.2345     0.0230     0.0002    0.0002   9461.6627    0.9999  ⋯
        β[5]    0.0209     0.0268     0.0003    0.0003   8849.2239    0.9999  ⋯
        β[6]   -0.0469     0.0186     0.0002    0.0002  10050.3263    1.0004  ⋯
        β[7]    0.0298     0.0189     0.0002    0.0002   9393.7048    0.9999  ⋯
        β[8]   -0.0370     0.0285     0.0003    0.0002   9875.2924    0.9999  ⋯
        β[9]    0.0301     0.0139     0.0001    0.0001   9648.7479    1.0001  ⋯
                                                            1 column omitted

Quantiles
  parameters      2.5%     25.0%     50.0%     75.0%     97.5%
      Symbol   Float64   Float64   Float64   Float64   Float64

           λ    0.0854    0.1249    0.1549    0.1937    0.3100
           ν    0.7862    2.2735    4.5396   11.1572  147.0323
        β[1]   -0.5099   -0.2325   -0.1163   -0.0225    0.1383
        β[2]   -0.0639   -0.0461   -0.0369   -0.0280   -0.0108
        β[3]   -0.0528   -0.0354   -0.0266   -0.0174   -0.0001
        β[4]    0.1899    0.2190    0.2343    0.2499    0.2793
        β[5]   -0.0311    0.0029    0.0207    0.0388    0.0746
        β[6]   -0.0836   -0.0597   -0.0469   -0.0341   -0.0107
        β[7]   -0.0068    0.0169    0.0297    0.0427    0.0668
        β[8]   -0.0930   -0.0563   -0.0367   -0.0175    0.0191
        β[9]    0.0026    0.0208    0.0302    0.0394    0.0573
```

Bayesian Poisson Regression with Horse Shoe Prior

In [8]:
```
mod_HS = fit(@formula(FTHG ~ HS + AS + HST + AST + HC + AC +B365H + B365A)
                  ,df_train,PoissonRegression()
                  ,Prior_HorseShoe())
```

```
┌ Info: Found initial step size
└   ϵ = 3.4694469519536144e-19
Sampling: 100%|█████████████████████████████████████████| Time: 0:01:46
```

Out[8]:
```
Formula: FTHG ~ 1 + HS + AS + HST + AST + HC + AC + B365H + B365A
Link: CRRao.Identity(CRRao.Identity_Link)
Chain: Chains MCMC chain (1000×31×1 Array{Float64, 3}):

Iterations        = 501:1:1500
Number of chains  = 1
Samples per chain = 1000
Wall duration     = 110.14 seconds
Compute duration  = 110.14 seconds
parameters        = τ, λ[1], λ[2], λ[3], λ[4], λ[5], λ[6], λ[7], λ[8], λ[9], β[1], β[2], β[3], β[4], β[5], β
[6], β[7], β[8], β[9]
internals         = lp, n_steps, is_accept, acceptance_rate, log_density, hamiltonian_energy, hamiltonian_ener
gy_error, max_hamiltonian_energy_error, tree_depth, numerical_error, step_size, nom_step_size

Summary Statistics
  parameters      mean       std   naive_se      mcse        ess      rhat   e ⋯
      Symbol   Float64   Float64    Float64   Float64    Float64   Float64     ⋯

           τ    0.0635    0.0458     0.0014    0.0058    44.9252    0.9994     ⋯
        λ[1]    2.4507    3.7919     0.1199    0.5050    31.0019    1.0056     ⋯
        λ[2]    2.8976    3.6569     0.1156    0.5497    21.7962    1.0218     ⋯
        λ[3]    0.9336    1.0848     0.0343    0.1025    86.6476    1.0091     ⋯
        λ[4]   10.0035   18.6598     0.5901    2.4316    49.6860    1.0259     ⋯
        λ[5]    0.8682    0.9239     0.0292    0.1033    73.0142    1.0010     ⋯
        λ[6]    2.4317   10.2987     0.3257    0.6058   216.5385    1.0014     ⋯
        λ[7]    1.0051    1.6633     0.0526    0.2303    44.5547    1.0030     ⋯
        λ[8]    1.3956    1.7142     0.0542    0.1553    62.8062    1.0079     ⋯
        λ[9]    0.7882    0.7374     0.0233    0.0851    59.2560    1.0015     ⋯
        β[1]   -0.0808    0.1491     0.0047    0.0146    46.7113    1.0103     ⋯
        β[2]   -0.0369    0.0147     0.0005    0.0011    85.9806    1.0128     ⋯
        β[3]   -0.0189    0.0125     0.0004    0.0009   181.2878    0.9990     ⋯
        β[4]    0.2339    0.0233     0.0007    0.0013   212.4691    1.0107     ⋯
        β[5]    0.0042    0.0200     0.0006    0.0011   378.1641    0.9999     ⋯
        β[6]   -0.0414    0.0200     0.0006    0.0014   111.3421    1.0055     ⋯
        β[7]    0.0148    0.0174     0.0006    0.0010   241.5746    1.0093     ⋯
        β[8]   -0.0292    0.0275     0.0009    0.0025    73.2588    1.0097     ⋯
        β[9]    0.0207    0.0146     0.0005    0.0013    75.8727    1.0008     ⋯
                                                            1 column omitted

Quantiles
  parameters      2.5%     25.0%     50.0%     75.0%     97.5%
      Symbol   Float64   Float64   Float64   Float64   Float64

           τ    0.0137    0.0325    0.0507    0.0776    0.2020
        λ[1]    0.1064    0.4433    1.1526    2.4544   14.6567
        λ[2]    0.0976    0.6809    1.3315    3.6645   12.9160
        λ[3]    0.0448    0.3038    0.5982    1.1379    4.2300
        λ[4]    0.9601    2.6341    4.7961    9.3348   63.7628
        λ[5]    0.0856    0.3238    0.6016    1.0793    3.3642
        λ[6]    0.1319    0.6039    1.0761    1.9451   10.7852
        λ[7]    0.1014    0.3468    0.6334    1.0848    3.3018
        λ[8]    0.0847    0.4637    0.8732    1.6833    5.2768
        λ[9]    0.0463    0.3072    0.5716    1.0311    2.7116
        β[1]   -0.4956   -0.1148   -0.0197    0.0042    0.0765
        β[2]   -0.0657   -0.0465   -0.0364   -0.0273   -0.0072
        β[3]   -0.0439   -0.0273   -0.0183   -0.0104    0.0030
        β[4]    0.1911    0.2182    0.2325    0.2492    0.2817
        β[5]   -0.0326   -0.0072    0.0025    0.0136    0.0492
        β[6]   -0.0776   -0.0558   -0.0426   -0.0284   -0.0011
        β[7]   -0.0143    0.0019    0.0127    0.0258    0.0529
        β[8]   -0.0869   -0.0479   -0.0257   -0.0071    0.0121
        β[9]   -0.0030    0.0096    0.0209    0.0303    0.0500
```

In [9]:
```
mod_HS_small = fit(@formula(FTHG ~ HS + HST + HC)
                   ,df_train,PoissonRegression()
                   ,Prior_HorseShoe())
```

```
┌ Info: Found initial step size
└   ϵ = 0.0001953125
Sampling: 100%|█████████████████████████████████████████| Time: 0:00:03
```

Out[9]:
```
Formula: FTHG ~ 1 + HS + HST + HC
Link: CRRao.Identity(CRRao.Identity_Link)
Chain: Chains MCMC chain (1000×21×1 Array{Float64, 3}):

Iterations        = 501:1:1500
Number of chains  = 1
Samples per chain = 1000
Wall duration     = 6.08 seconds
Compute duration  = 6.08 seconds
parameters        = τ, λ[1], λ[2], λ[3], λ[4], β[1], β[2], β[3], β[4]
internals         = lp, n_steps, is_accept, acceptance_rate, log_density, hamiltonian_energy, hamiltonian_ener
gy_error, max_hamiltonian_energy_error, tree_depth, numerical_error, step_size, nom_step_size

Summary Statistics
  parameters      mean       std   naive_se      mcse        ess      rhat   e ⋯
      Symbol   Float64   Float64    Float64   Float64    Float64   Float64     ⋯

           τ    0.3946    0.3513     0.0111    0.0188   368.5988    1.0111     ⋯
        λ[1]    2.5646    4.2298     0.1338    0.1868   468.5260    1.0099     ⋯
        λ[2]    0.7144    1.3416     0.0424    0.0436   804.4869    0.9990     ⋯
        λ[3]    1.9189    2.8678     0.0907    0.1137   542.7584    1.0009     ⋯
        λ[4]    0.7556    2.0590     0.0651    0.0766   765.0860    1.0020     ⋯
        β[1]   -0.3863    0.1293     0.0041    0.0062   524.4397    1.0021     ⋯
        β[2]   -0.0284    0.0148     0.0005    0.0008   341.0983    1.0004     ⋯
        β[3]    0.2431    0.0221     0.0007    0.0012   379.9273    0.9998     ⋯
        β[4]   -0.0365    0.0179     0.0006    0.0009   559.1591    1.0002     ⋯
                                                               1 column omitted

Quantiles
  parameters      2.5%     25.0%     50.0%     75.0%     97.5%
      Symbol   Float64   Float64   Float64   Float64   Float64

           τ    0.0565    0.1733    0.2995    0.4893    1.3802
        λ[1]    0.2632    0.7968    1.4737    2.7810   10.9341
        λ[2]    0.0296    0.1608    0.3478    0.7931    3.3743
        λ[3]    0.2386    0.6768    1.1526    1.9833    8.9125
        λ[4]    0.0382    0.1736    0.4038    0.7734    3.3412
        β[1]   -0.6304   -0.4765   -0.3880   -0.2982   -0.1319
        β[2]   -0.0576   -0.0384   -0.0284   -0.0186    0.0001
        β[3]    0.2012    0.2279    0.2433    0.2578    0.2873
        β[4]   -0.0710   -0.0484   -0.0356   -0.0237   -0.0038
```

In [10]:
```
pred_FTHG = predict(mod_HS_small,df_train);
```

In [11]:
```
pred_FTHG
```

Out[11]:
```
380-element Vector{Float64}:
 0.8924489238533787
 0.953105286404616
 1.1377269787383852
 0.6915580613565018
 0.6615862087919823
 1.4811993289246943
 0.844390692927946
 0.560038291818685
 1.5504011585488242
 2.343971193196086
 0.8991012783133118
 0.8987351119621247
 2.247688053598424
 ⋮
 1.0887707750031057
 0.9727343554643355
 0.9816409372101816
 1.105893332170171
 0.3846014710296586
 2.803581869191904
 1.6514310571553237
 0.8137947180728899
 4.226107162034165 5
 0.7518355392578444
 2.3325122405509036
 0.9727343554643355
```

In [15]:
```
cor(pred_FTHG,df_train.FTHG)^2
# Hence Bayesian regression gives better predictive power
```

Out[15]:
```
0.4270550503539198
```

In [18]:
```
sqrt(mean((pred_FTHG-df_train.FTHG).^2))
```

Out[18]:
```
1.000902177196898
```

In [23]:
```
using Plots
scatter(pred_FTHG,df_train.FTHG,label="")
xlabel!("predicted goal by Home Team")
ylabel!("actual number of goals by Home Team")
```

Out[23]: