UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2016:46

Bivariate Models to Predict Football Results

Joel Lidén

Examensarbete i matematik, 15 hp
Handledare: Rolf Larsson
Examinator: Jörgen Östensson
December 2016

Department of Mathematics
Uppsala University

Bivariate Models to Predict Football Results

Joel Lidén
Degree Project C in Mathematics
Uppsala University
Supervisor: Rolf Larsson

Autumn 2016

December 5, 2016

Contents

1 Abstract **3**

2 Introduction **3**

3 Seasonal Data from European Football Leagues **4**

4 Theory **5**

4.1 The Naive Model 5

4.2 Poisson Regression Estimation 7

4.3 Negative Binomial Regression Estimation 9

4.4 Deviance Goodness-of-fit 11

4.5 Overdispersion in a GLM 12

4.6 Using a Discrete Copula 13

4.7 Arbitrage Strategy 16

4.8 The Evaluation Program 17

5 Results **19**

5.1 Poisson Distribution Assumption 19

5.2 Negative Binomial Distribution Assumption 22

5.3 Goodness-of-fit Test 25

5.4 Independence Assumption 26

5.5 Fitting the Poisson Model 27

5.6 Fitting the Negative Binomial Model 38

5.7 Results of Betting Evaluations 41

6 Discussion **47**

7 References **49**

8 Appendix **50**

8.1 R Code 50

8.2 Evaluation Plots 53

1 Abstract

In this paper different models predicting full-time scores of football games will be implemented and tested using historical data. Models using a bivariate distribution for number of home goals and away goals will be fitted and tested in practice. Profitability against several bookmakers will be investigated using evaluations. The models will also be tested against random betting, to see how they compete with both the bookmakers and pure chance. Evaluations and statistical tests will be carried out using the R software.

2 Introduction

Sports betting has a long tradition and history, with football betting being a multi billion dollar industry. Today, with the impact of online betting services, it's easier than ever to place a bet. In a standard game, a bettor can choose whether to bet on the home team winning, the away team winning or a draw. There are also other types of bets such as Asian handicap, exact results, number of goals being scored etc. However, in this paper, only the standard types of bets will be considered, i.e. home win, away win or a draw. Since many bettors may have a bias towards their favorite team winning, or betting with their "gut feeling", many bets are not objectively considered, and have a negative expected profit in the long run. The sports betting companies also have an edge for each game (usually between 2-8 %) which is their profit margin. Since it's virtually impossible to predict probabilities of a football game exactly, it is possible to profit from football betting in the long run, even though it's quite difficult. The odds also vary slightly between different bookmakers, which is why it's wise to be able to use multiple bookmakers, in order to always get the best possible odds.

Today, vast amounts of data displaying historical football results are available for download completely for free. Statistical softwares such as R allow for analysis and model building using these data. Evaluations can also be made since the odds of different bookmakers are listed for each game played, and conclusions can be drawn whether a model is profitable in the long run or not. So therefore, there are more opportunities than ever to dig deep into the data and use statistical tools to predict a winner. In this paper, the goal is to find a statistical model accurate enough to be a consistent winner in the long run.

3 Seasonal Data from European Football Leagues

The data being used for evaluations and statistical tests includes historical results from 22 major European leagues as far back as the season 1993/1994 [2]. Included in the data are also odds from several bookmakers for each game played, which makes it possible to run evaluations using a prediction model and see if the model would have been profitable or not. Since leagues differ in terms of aggressiveness and defensiveness, it's interesting to see how many goals are being scored on average for home and away team and also if some leagues seem to have higher variance than others. Looking at the 2015/2016 season, we can see averages and variances for different leagues in Table 1.

Country	League name	Average home goals	Average away goals	Variance home goals	Variance away goals
Belgium	Pro League	1.68	1.22	1.64	1.12
Germany	Bundesliga	1.57	1.26	1.86	1.28
Germany	Bundesliga 2	1.4	1.25	1.54	1.35
England	Premier League	1.49	1.21	1.59	1.32
England	Championship	1.35	1.07	1.31	1.07
England	League 1	1.45	1.19	1.55	1.3
England	League 2	1.38	1.29	1.3	1.32
England	Conference	1.46	1.18	1.58	1.24
France	Ligue 1	1.44	1.09	1.64	1.32
France	Ligue 2	1.3	1	1.33	1.01
Greece	Super League	1.32	0.94	1.57	1.08
Italy	Serie A	1.47	1.11	1.49	1.17
Italy	Serie B	1.47	0.97	1.39	1
Netherlands	Eredivisie	1.63	1.35	1.67	1.55
Portugal	Primeira Liga	1.51	1.2	1.8	1.42
Scotland	Premier League	1.5	1.35	1.85	1.42
Scotland	Championship	1.49	1.19	1.58	1.42
Scotland	League One	1.62	1.37	2.2	1.52
Scotland	League Two	1.56	1.31	1.74	1.59
Spain	La Liga	1.62	1.13	2.1	1.32
Spain	Segunda Liga	1.34	0.92	1.21	0.94
Turkey	Super League	1.55	1.16	1.68	1.21

Table 1: Averages and variances for 22 European leagues during the season 2015/2016

By looking at the data in Table 1, it can be seen that the averages and variances are fairly equal to one another, which would be the case if the number of home goals and number of away goals are indeed Poisson distributed. However, in many leagues the variance is slightly larger than the mean, which is a sign of overdispersion in the Poisson case. The Poisson distribution assumption will be investigated further in section 5.1 and section 5.3.

4 Theory

4.1 The Naive Model

A common model that can be found in popular science and at sports betting webpages for predicting football results uses an attack- and a defense coefficient to estimate the expected number of goals for both the home team and the away team [1]. The model then assumes that both the number of home goals and away goals are two independent random variables that follow a marginal Poisson distribution. The total result in the game is then assumed to follow a bivariate Poisson distribution. Since the two random variables are assumed to be independent, the bivariate Poisson density will simply be the product of the two marginal Poisson densities. Mathematically, it can be expressed as follows:

$$X = \text{"Number of home goals"} \sim Poisson(\lambda_x)$$

$$Y = \text{"Number of away goals"} \sim Poisson(\lambda_y)$$

$$P(X = x) = \frac{\lambda_x^x e^{-\lambda_x}}{x!} \qquad\qquad P(Y = y) = \frac{\lambda_y^y e^{-\lambda_y}}{y!}$$

$$P(X = x, Y = y) = P(X = x)P(Y = y) = \frac{\lambda_x^x \lambda_y^y e^{-(\lambda_x + \lambda_y)}}{x!y!}$$

The expectations for each of the two random variables (i.e. λ_x and λ_y) are then calculated by assigning an attack and a defense coefficient for both the home and away team. The attack coefficient for the home team is estimated by taking the ratio of the average number of home goals scored by the home team this season and the total average of all home goals scored this season. The attack coefficient for the away team is then estimated in an analogous manner, but by using the ratio of the average number of away goals scored by the away team and the total average of all away goals scored this season.

The defense coefficient for the home team is then calculated by taking the ratio of the average number of goals conceded by the home team at home and the total average of goals conceded at home this season. And finally, the defense coefficient for the away team is calculated by taking the ratio of the average number of goals conceded away for the away team and the total average of goals conceded away this season.

Now, having calculated both attack and defense coefficients for both the home team and the away team, the expected number of goals for the home team (i.e. λ_x) is simply the attack coefficient for the home team multiplied by the defense coefficient of the away team multiplied by the total average number of goals scored at home this season. The expected number of away goals is simply the attack coefficient of the away team multiplied by the defense coefficient of the home team multiplied by the total average number of goals scored away this season. Estimating attack and defense coefficients as well as expectations can be expressed as follows:

$$\text{Attack}_x = \frac{\text{Average n.o. homegoals scored by the home team}}{\text{Average n.o. homegoals scored in total this season}}$$

$$\text{Defense}_x = \frac{\text{Average n.o. goals conceded by the home team at home}}{\text{Average n.o. goals conceded at home in total this season}}$$

$$\text{Attack}_y = \frac{\text{Average n.o. awaygoals scored by the away team}}{\text{Average n.o. awaygoals scored in total this season}}$$

$$\text{Defense}_y = \frac{\text{Average n.o. goals conceded by the away team playing away}}{\text{Average n.o. goals conceded away in total this season}}$$

$$E(X) = \lambda_x = \text{Attack}_x * \text{Defense}_y * \text{Average n.o. homegoals scored in total this season}$$

$$E(Y) = \lambda_y = \text{Attack}_y * \text{Defense}_x * \text{Average n.o. awaygoals scored in total this season}$$

Since the average number of home goals scored by the home team during the season must be the same as the average number of goals conceded by the away team during the season, the denominator of the home attack coefficient must be the same as the denominator of the away defense coefficient. In an analogous manner, the denominator of the home defense coefficient must be the same as the denominator of the away attack coefficient. Now, having expectations for both random variables, the bivariate Poisson distribution can be used to calculate each full time result, by using the bivariate density function. Calculating probabilities for a home win, away win, and a draw is then easily done by summing up the probabilities of the different results. While running evaluations, all probabilities of results from 0-0 to 9-9 (i.e. 100 different outcomes) have been calculated. The probability of a team scoring 10 or more goals in a game against an opponent in the same league is so small that it can be considered negligible.

4.2 Poisson Regression Estimation

A Generalized Linear Model generalizes ordinary linear regression by relating the response variable to the linear model by a link function. This relaxes the normality assumption in the ordinary linear regression case, so that the response variable can follow other distributions than the normal. Therefore, count data such as the number of goals for home and away team can be modelled using Poisson regression, where the response variable is Poisson distributed. Since the number of goals for home and away team can be approximately regarded as Poisson, this approach is reasonable. In the Poisson case, the link function must first be identified by rewriting the probability mass function as an exponential dispersion family. The general form of an exponential dispersion family, where N observations $(y_1, y_2, ..., y_N)$ of the response variable Y, with probability mass or density function $f(y_i; \theta_i, \phi)$ for y_i, can be expressed as [3]:

$$f(y_i; \theta_i; \phi) = \exp\left(\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi)\right)$$

In the exponential dispersion family, θ_i is the canonical parameter which depends on a model of linear predictors. The dispersion parameter is noted ϕ, and is often known. Using the probability mass function of the Poisson distribution, the exponential dispersion form can be retrieved as follows:

$$f(y_i; \theta_i; \phi) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!} = \exp\Big(\log(e^{-\lambda_i}) + \log(\lambda_i^{y_i}) - \log(y_i!)\Big)$$

$$\Rightarrow f(y_i; \theta_i; \phi) = \exp\left(y_i \log(\lambda_i) - \lambda_i - \log(y_i!)\right)$$

$$\Rightarrow a(\phi) = 1 \quad \theta_i = \log(\lambda_i) \Leftrightarrow \lambda_i = e^{\theta_i} \quad b(\theta_i) = \lambda_i = e^{\theta_i} \quad c(y_i, \phi) = -\log(y_i!) = \log(\frac{1}{y_i!})$$

In general, the expectation and variance of Y_i can be expressed as follows: [3]

$$E(Y_i) = \mu_i = b'(\theta_i) \qquad\qquad Var(Y_i) = b''(\theta_i)a(\phi)$$

Since $b(\theta_i) = \lambda_i$ and $a(\phi) = 1$ in the Poisson case, the expectation and variance are computed as follows:

$$E(Y_i) = b'(\theta_i) = e^{\theta_i} = \lambda_i \qquad Var(Y_i) = b''(\theta_i)a(\phi) = e^{\theta_i} = \lambda_i$$

So therefore, the expectation is equal to the variance as expected in the Poisson case. The log-link η_i can be identified as $\eta_i = g(\lambda_i) = \theta_i = \log(\lambda_i)$. So therefore the log of the expectation, i.e. $\log(\lambda_i)$, can be modelled by the predictors x_{ij}, where β_j is the estimated coefficient for each predictor. This can be expressed as follows:

$$\eta_i = \log(\lambda_i) = \sum_k \beta_k x_{ik},\ i = 1, ..., N$$

The scoring intensities for the home team and the away team (i.e. λ_x and λ_y) can therefore be modelled by Poisson regression, keeping the attack- and defense coefficients from 4.1 as predictors. Using the log-link in the GLM, the scoring intensities for the home team i versus the away team j can be modelled as follows:

$$\log(\lambda_{x,i}) = \alpha + \beta_1 attack_{x,i} + \beta_2 defense_{y,j}$$

$$\log(\lambda_{y,j}) = \alpha + \beta_1 attack_{y,j} + \beta_2 defense_{x,i}$$

4.3 Negative Binomial Regression Estimation

Just like the Poisson case, a negative binomial GLM can be fitted to count data such as the number of goals for the home and away team. Since the Poisson case is limited to having equal expectation and mean, a negative binomial GLM allows for overdispersion, i.e. larger variance than the mean. The goodness-of-fit test was also accepted for an approximation using the negative binomial distribution. Therefore, a negative binomial GLM might prove to be a better model than the Poisson GLM. A link function can also be found for the negative binomial case. Starting with the probability mass function to retrieve the exponential dispersion form, and identify the link function, the following steps are taken:

$$f(y_i;\theta_i;\phi) = \frac{\Gamma(y_i+k)}{\Gamma(k)\Gamma(y_i+1)}\left(\frac{k}{\mu_i+k}\right)^k\left(\frac{\mu_i}{\mu_i+k}\right)^{y_i}$$

$$\Rightarrow f(y_i;\theta_i;\phi) = \exp\left(\log\left(\left(\frac{\mu_i}{\mu_i+k}\right)^{y_i}\right)+\log\left(\left(\frac{k}{\mu_i+k}\right)^{k}\right)+\log\left(\frac{\Gamma(y_i+k)}{\Gamma(k)\Gamma(y_i+1)}\right)\right)$$

$$\Rightarrow f(y_i;\theta_i;\phi) = \exp\left(y_i\log\left(\frac{\mu_i}{\mu_i+k}\right)+k\log\left(\frac{k}{\mu_i+k}\right)+\log\left(\frac{\Gamma(y_i+k)}{\Gamma(k)\Gamma(y_i+1)}\right)\right)$$

$$\Rightarrow a(\phi)=1 \quad \theta_i=\log\left(\frac{\mu_i}{\mu_i+k}\right) \quad b(\theta_i)=-k\log\left(\frac{k}{\mu_i+k}\right) \quad c(y_i,\phi)=\log\left(\frac{\Gamma(y_i+k)}{\Gamma(k)\Gamma(y_i+1)}\right)$$

The expectation and variance in the negative binomial case can therefore be retrieved according to:

$$\theta_i=\log\left(\frac{\mu_i}{\mu_i+k}\right)\Leftrightarrow\frac{\mu_i}{\mu_i+k}=e^{\theta_i}\Rightarrow\mu_i=\frac{ke^{\theta_i}}{1-e^{\theta_i}}$$

$$b(\theta_i)=-k\log\left(\frac{k}{\mu_i+k}\right)=k\log\left(\frac{\mu_i+k}{k}\right)=k\log\left(\frac{\frac{ke^{\theta_i}}{1-e^{\theta_i}}+k}{k}\right)=k\log\left(\frac{1}{1-e^{\theta_i}}\right)=-k\log(1-e^{\theta_i})$$

$$E(Y_i)=b'(\theta_i)=\frac{ke^{\theta_i}}{1-e^{\theta_i}}=\frac{\left(\frac{k\mu_i}{\mu_i+k}\right)}{\left(\frac{k}{\mu_i+k}\right)}=\mu_i$$

$$Var(Y_i)=b''(\theta_i)a(\phi)=\frac{ke^{\theta_i}}{(1-e^{\theta_i})^2}=\frac{\frac{k\mu_i}{\mu_i+k}}{\left(1-\frac{\mu_i}{\mu_i+k}\right)^2}=\mu_i+\frac{1}{k}\mu_i^2$$

As can be seen, $Var(Y_i) \longrightarrow \mu_i$ when $k \longrightarrow \infty$. So the negative binomial distribution converges to Poisson in this case [3]. Therefore, the Poisson is in fact a special case of the negative binomial distribution. Now, the canonical link function in the negative binomial case can be identified as $\eta_i = g(\mu_i) = \theta_i = \log\left(\frac{\mu_i}{\mu_i+k}\right) = \mathbf{x_i}\boldsymbol{\beta}$. Since $\mu_i > 0$, the image of $g(\mu_i) \in (-\infty, 0)$. Therefore the canonical link is not suitable since the linear prediction model must be able to take on positive values. So a better choice is to use a log link similar to the Poisson model. In that case, the log of the expectation can be modelled in the same way, i.e.:

$$\log(\mu_{x,i}) = \alpha + \beta_1 attack_{x,i} + \beta_2 defense_{y,j}$$

$$\log(\mu_{y,j}) = \alpha + \beta_1 attack_{y,j} + \beta_2 defense_{x,i}$$

4.4 Deviance Goodness-of-fit

When a Generalized Linear Model has been fitted, a deviance goodness-of-fit test is a good way to assess the explanatory power of the model. In such a test, the actual model is being compared to a saturated GLM which has a shape parameter for each observation, yielding a perfect fit. This model sounds like an excellent choice, but in reality it does not smooth the data and it does not have the advantages that a simpler model has, which uses only a few predictors. It is, however, useful for testing the fit of other models. The following test statistic is being used, where $L(\hat{\mu};\mathbf{y})$ is the maximized log likelihood for the model being tested, and $L(\mathbf{y};\mathbf{y})$ is the maximized log likelihood in the saturated case [3]:

$$-2[L(\hat{\mu};\mathbf{y})-L(\mathbf{y};\mathbf{y})]=2\sum_i[y_i\tilde{\theta}_i-b(\tilde{\theta}_i)]/a(\phi)-2\sum_i[y_i\hat{\theta}_i-b(\hat{\theta}_i)]/a(\phi)$$

$$2\sum_i\omega_i[y_i(\tilde{\theta}_i-\hat{\theta}_i)-b(\tilde{\theta}_i)+b(\hat{\theta}_i)]/\phi=D(\mathbf{y};\hat{\mu})\sim\chi^2_{n-p}$$

The last two equalities hold since $a(\phi)=\frac{\phi}{\omega_i}=1$. The last expression $D(\mathbf{y};\hat{\mu})$ is called the scaled deviance and has an approximate chi-squared distribution with $n-p$ degrees of freedom, where n is equal to the number of parameters in the saturated model and p is the number of parameters in the model being tested. If the model is a good fit for the data, the deviance will be small, since the observed values are close to the predicted ones given by the model. Testing the null hypothesis that the proposed model is correctly specified means that a large p-value indicates a good fit. The deviance for Poisson and negative binomial GLM's can be expressed as follows:

$$D(\mathbf{y};\hat{\mu})_{\text{Poisson}}=2\sum_i[y_i\log\left(\frac{y_i}{\hat{\mu}_i}\right)-y_i+\hat{\mu}_i]$$

$$D(\mathbf{y};\hat{\mu})_{\text{NegBin}}=2\sum_i y_i\left[\log\left(\frac{y_i(\hat{\mu}_i+k)}{\hat{\mu}_i(y_i+k)}\right)+k\log\left(\frac{\hat{\mu}_i+k}{y_i+k}\right)\right]$$

4.5 Overdispersion in a GLM

Overdispersion in a Generalized Linear Model is a sign of the variance being larger than the mean. In that case a negative binomial model is preferable over a Poisson model, since the Poisson assumption of the expectation being equal to the variance is clearly not the case. In a Poisson GLM, the estimated variance can be expressed as follows:

$$Var(Y_i) = \phi\mu_i$$

So the Poisson assumption indicates $\phi = 1$, which yields that the variance is equal to the expectation. If $\hat{\phi} > 1$ there is overdispersion in the model, and if $\hat{\phi} < 1$, there is underdispersion, where the variance is smaller than the expectation. To assess whether or not a model shows evidence of overdispersion, the ratio of the residual deviance over the degrees of freedom can be evaluated. The residual deviance is simply the sum of all deviance residuals, and can be expressed as follows:

$$D(\mathbf{y}, \hat{\mu}) = \sum_i d_i$$

$$d_i = 2\omega_i[y_i(\tilde{\theta}_i - \hat{\theta}_i) - b(\tilde{\theta}_i) + b(\hat{\theta}_i)]$$

$$\hat{\phi} = \frac{\text{Residual deviance}}{\text{Degrees of freedom}} = \frac{D(\mathbf{y}, \hat{\mu})}{n-p} = \frac{\sum_i 2\omega_i[y_i(\tilde{\theta}_i - \hat{\theta}_i) - b(\tilde{\theta}_i) + b(\hat{\theta}_i)]}{n-p}$$

In the expression above, $n - p$ is the difference between the number of parameters in the saturated model, explained in section 4.4. If $\hat{\phi} > 1$, the Poisson assumption is not adequate, since the variance is larger than the expectation. In this case, a negative binomial GLM is preferable, since it allows for overdispersion. As discussed in section 4.3, the variance of Y_i in a negative binomial GLM can be expressed as:

$$Var(Y_i) = \mu_i + \frac{1}{k}\mu_i^2$$

So therefore $Var(Y_i) = \mu_i + \frac{1}{k}\mu_i^2 \longrightarrow \mu_i$ as $k \longrightarrow \infty$, where k is the determined number of failures in the negative binomial probability mass function. In a negative binomial GLM, $k = \theta$, so if there is no overdisperion (or very little), θ will be estimated to be very large, since the variance converges to the μ_i, i.e. the variance in the Poisson case.

4.6 Using a Discrete Copula

When two discrete random variables are dependent, their joint cumulative distribution function can be found using a two-dimensional copula. The copula is therefore basically a linking function which links the two univariate margins together, and also allows for a dependence parameter θ. In this case, the two univariate margins are standard uniform, so mathematically the copula function can be expressed as follows [5]:

$$C(u_1, u_2|\theta) = P(U_1 \leqslant u_1, U_2 \leqslant u_2)$$

In this expression, U_1 and U_2 are independent and identically distributed standard uniform, i.e. $U[0, 1]$, and θ is a dependence parameter. If Y_j has a continuous cdf F_j, the transform $F_j(Y_j)$ must be uniformly distributed. Therefore, the joint bivariate cdf with marginal cdf's F_1 and F_2 can be expressed as follows [5]:

$$F(y_1, y_2) = P[Y_1 \leqslant y_1, Y_2 \leqslant y_2] = P[F_1(Y_1) \leqslant F_1(y_1), F_2(Y_2) \leqslant F_2(y_2)]$$

$$= P[U_1 \leqslant F_1(y_1), U_2 \leqslant F_2(y_2)] = C(F_1(y_1), F_2(y_2)|\theta)$$

An alternative way to express the copula is as follows:

$$F(y_1, y_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)) = C(u_1, u_2|\theta)$$

However, the assumption that Y_j has a continuous cdf does not hold true. Since both the Poisson and the negative binomial distribution are discrete, their marginal cumulative distribution functions are step functions with jumps at integer values, meaning that the inverse F_j^{-1} is not unique. So therefore a convention must be imposed, where usually the minimum of the interval is chosen. There are several different copula types, which have different domains of the dependence parameter θ. For constructing a bivariate cdf of Poisson and negative binomial marginals, a Frank copula type has been chosen since the domain $\theta \in (-\infty, \infty)$, i.e. the entire real line. So the Frank copula is a good choice in this case since it's very general, and can be expressed as follows [6]:

$$C(F_X(s), F_Y(t)) = \frac{1}{\theta} \log \left(1 + \frac{\exp(\theta F_X(s) - 1)(\exp(\theta F_Y(t)) - 1)}{\exp(\theta) - 1} \right) \quad (1)$$

In this case, $s, t = 0, 1, 2, ...$, i.e. the number of goals scored for home and away team. In the Poisson case, the cdf's for the number of home goals (X) and away goals (Y) can be expressed as follows:

$$u_X = F_X(s) = P(X \leqslant s) = \sum_{m=0}^{s} \frac{\lambda_x^m e^{-\lambda_x}}{m!} \quad u_Y = F_Y(t) = P(Y \leqslant t) = \sum_{l=0}^{t} \frac{\lambda_y^l e^{-\lambda_y}}{l!}$$

In the negative binomial case, where k is the fixed number of failures and the indices of the sums (m and l) are the number of successes, the cdf's can be expressed as follows:

$$u_X = F_X(s) = P(X \leqslant s) = \sum_{m=0}^{s} \frac{\Gamma(m+k)}{\Gamma(k)\Gamma(m+1)} \left(\frac{k}{\mu+k}\right)^k \left(\frac{\mu}{\mu+k}\right)^m$$

$$u_Y = F_Y(t) = P(Y \leqslant t) = \sum_{l=0}^{t} \frac{\Gamma(l+k)}{\Gamma(k)\Gamma(l+1)} \left(\frac{k}{\mu+k}\right)^k \left(\frac{\mu}{\mu+k}\right)^l$$

So therefore, u_X and u_Y can simply be plugged into the copula function to retrieve the probability of a certain result iteratively. Since the copula function is actually the joint cdf and not the joint pmf, probabilities of certain results can be computed as follows [6]:

$$P(X = 0, Y = 0) = C(F_X(0), F_Y(0))$$

$$P(X = s, Y = 0) = C(F_X(s), F_Y(0)) - C(F_X(s-1), F_Y(0)) \quad s = 1, 2, ...$$

$$P(X = 0, Y = t) = C(F_X(0), F_Y(t)) - C(F_X(0), F_Y(t-1)) \quad t = 1, 2, ...$$

$$P(X = s, Y = t) = C(F_X(s), F_Y(t)) - C(F_X(s-1), F_Y(t)) - C(F_X(s), F_Y(t-1))$$

$$+\mathrm{C}(\mathrm{F}_X(s-1), F_Y(t-1)) \quad s, t = 1, 2, ...$$

The dependence parameter θ allows for negative correlation, which is appropriate, since the sample data suggests negative correlation between home goals and away goals. Moreover, the dependence parameter in this case is not the Pearson correlation itself which lies in the interval $[-1, 1]$, but instead θ is converted into a measure of concordance such as the Kendall's τ. For the Frank copula, Kendall's τ can be expressed as follows [5]:

$$\tau = f(\theta) = 1 + \frac{4}{\theta}\left[\int_0^{\theta} \frac{a}{\theta(e^a - 1)} da - 1\right] \qquad (2)$$

So in this case $\tau = f(\theta) \Leftrightarrow \theta = f^{-1}(\tau)$. Since the function f is invertible, θ can be estimated using an estimate of Kendall's τ. The R software can be used for both the estimate of Kendall's τ and f^{-1}. So therefore, an estimate of the dependence parameter θ can be found.

4.7 Arbitrage Strategy

A strategy that is guaranteed to make at least a small profit is by using arbitrage opportunities in the sportsbook market. Since different bookmakers are pricing the odds differently, there are sometimes opportunities where it's possible to bet on all three outcomes of a game (i.e. home win, draw and away win) and make a small but guaranteed profit. Such opportunities are called arbitrage opportunities. If O_{ij} is the maximum odds of outcome i for $i = 1, 2, 3$, in game j, there is an arbitrage opportunity in game j if the following inequality holds [7]:

$$\frac{1}{O_{1j}} + \frac{1}{O_{2j}} + \frac{1}{O_{3j}} < 1$$

As an arbitrary example, if the odds of a home win at a bookmaker A is 2, the odds of a draw at another bookmaker B is 3 and the odds of an away win at yet another bookmaker C is 8, the sum of the reciprocal of the odds would be:

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{8} = 0.958 < 1$$

The sum of the reciprocal of the odds is indeed smaller than 1, so there is an arbitrage opportunity in this game. If a standard bet size is for example 20 units, the bet size for each outcome can be chosen relative to the odds, so that a payout of 20 units is guaranteed regardless of whether the game ends in a home win, draw or an away win.

Outcome	Home win	Draw	Away win
Bookmaker	A	B	C
Odds	2	3	8
Betsize	10	6.67	2.5
Return	20	20	20
Total betsize	19.17	19.17	19.17
Profit	0.83	0.83	0.83

Table 2: Example depicting the outcome of an arbitrage opportunity

As can be seen in Table 2, a return of 20 units is guaranteed while risking only 19.17 units. Therefore a profit of 0.83 units is guaranteed regardless of how the game ends. By using this strategy, small profits can be made without taking any risks, although arbitrage opportunities are quite rare, and can only be found in a small portion of games.

4.8 The Evaluation Program

To test if any of the three models (the naive model, the Poisson model and the negative binomial model) can be profitable in the long run, evaluations using the R software can be made, since for every match played in the data set, the odds of different bookmakers are listed. The evaluation program selects prediction data which consists of historical matches from the same league as the match being played, where the selected matches are within a time frame of one year from the current match being played. This time frame can also be varied since having a small time frame causes fewer prediction matches, which means higher variability in the data selected. A larger time frame reduces variability, but instead might include data that is not relevant anymore since teams change from one season to another, and since the performance of a certain team is not constant over time.

When the prediction data has been selected, attack and defense coefficients for each team in the league are estimated. Then two Poisson and two negative binomial GLM's are fitted, using the number of home and away goals as response variables, with the attack and defense coefficients as predictors, as explained in section 4.2 and section 4.3. After that, an expectation of the number of goals made by the home team and the away team can be calculated by simply plugging in the attack and the defense coefficients of the respective teams in the fitted GLM's. Using the estimated expectations, the marginal cumulative distribution functions can be used to estimate the probabilities of the cumulative number of goals that each team will score. These cumulative probabilities can then simply be plugged into the copula function described in equation 1.

Since the copula function also uses a dependence parameter as an input variable, the dependence parameter can be estimated by first letting the R software calculate an estimate of the Kendall's τ correlation between home and away goals in the prediction data. Equation 2 expresses Kendall's τ as a function of the dependence parameter θ. Using the inverse of this function, an estimate of the dependence parameter, i.e. $\hat{\theta}$, can be found since the inverse uses Kendall's τ as the input variable. Now when all input variables of the copula function have been estimated, the copula function, which can be described as a bivariate cumulative distribution function, can be used to isolate estimated probabilities of certain results.

The naive model estimates the expectations of home and away goals by simply multiplying the attack and the defense coefficients and also the average number of goals being scored at home and away during the season, as described in section 4.1. Since the naive model also assumes independence, the marginal Poisson densities for home and away goals can be used directly to estimate the probability of a certain result, since the joint density is simply the product of the marginal densities if the random variables are independent. This is also described in section 4.1.

Now, since probabilities can be estimated for all three models, 10 by 10 matrices can be made for all three models, with estimated probabilities of 100 results ranging from 0-0 to 9-9 as entries. As stated earlier, the probability of a team scoring 10 goals or more against a team in the same league is so small that it can be considered negligible. The matrices use home goals as rows and away goals as columns, so the traces of the matrices become the estimated probabilities of a draw. In a similar fashion, the under triangular matrices yield probabilities of a home win, and the upper triangular matrices the probabilities of an away win.

The listed odds given by the bookmakers are European odds, where the odds is simply the reciprocal of the probability. So the estimated probabilities can therefore easily be converted to estimated odds. Then the actual odds and the estimated odds can be compared by the following ratio, where r is a given margin of error that can be varied:

$$\frac{\text{Maximum odds given by the bookmakers}}{\text{Calculated odds}} > 1 + r$$

In this case the evaluation program selects the largest odds given by the following bookmakers: Bet365, Bwin, Interwetten, Ladbrokes, Pinnacle Sports, William Hill and BetVictor. The program then selects to bet on either a home win, draw or an away win based on which of the three outcomes that yields the largest ratio.

For each match played, a random bet is also made. In this case, historical frequencies of home wins, draw and away wins are used to determine probabilities of each outcome. The evaluation program then simply bets on one of the three outcomes with the assigned probability. The maximum odds for a given bet is selected also in the random case, in order for the bets made by the models to be comparable to the random bets. Finally, the evaluation program also checks for arbitrage opportunities, described in section 4.7, in each game, and bets if there is one. Since arbitrage opportunities are quite rare, there won't be nearly as many arbitrage bets as bets made by the prediction models.

5 Results

5.1 Poisson Distribution Assumption

The naive model assumes that both marginal distributions are Poisson-distributed. This assumption can be tested using data of historical results from the 2015/2016 season. In Figure 1, one can see that the expected Poisson frequencies are quite close to the actual observed frequencies for number of goals scored by the home team. The expectation of the Poisson distribution (i.e. λ) is taken to be the overall average of the number of goals scored by the hometeam during the 2015/2016 season in the 22 different European leagues. In an analogous manner, a plot comparing the actual against the expected Poisson frequencies for number of goals scored away can be seen in Figure 2.

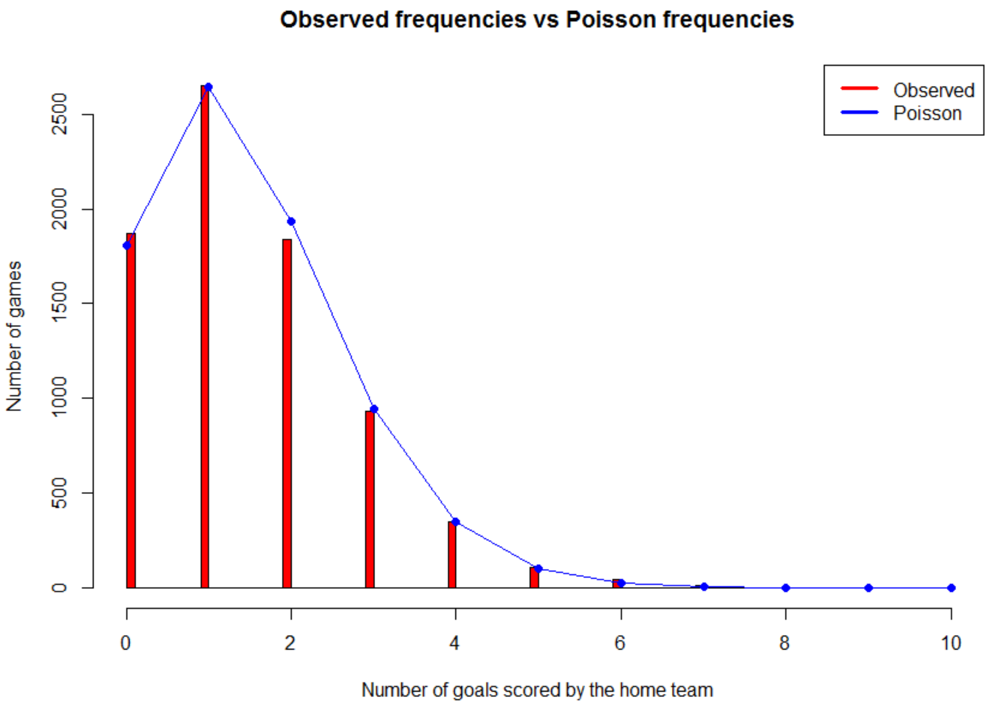


Figure 1: Plot comparing actual against expected Poisson frequencies for n.o. home goals

Home goals	Observed	Expected Poisson
0	1870	1806
1	2652	2645
2	1840	1936
3	933	945
4	349	346
5	107	101
6	45	25
7+	14	6

Table 3: Comparing observed against expected Poisson frequencies for n.o. home goals

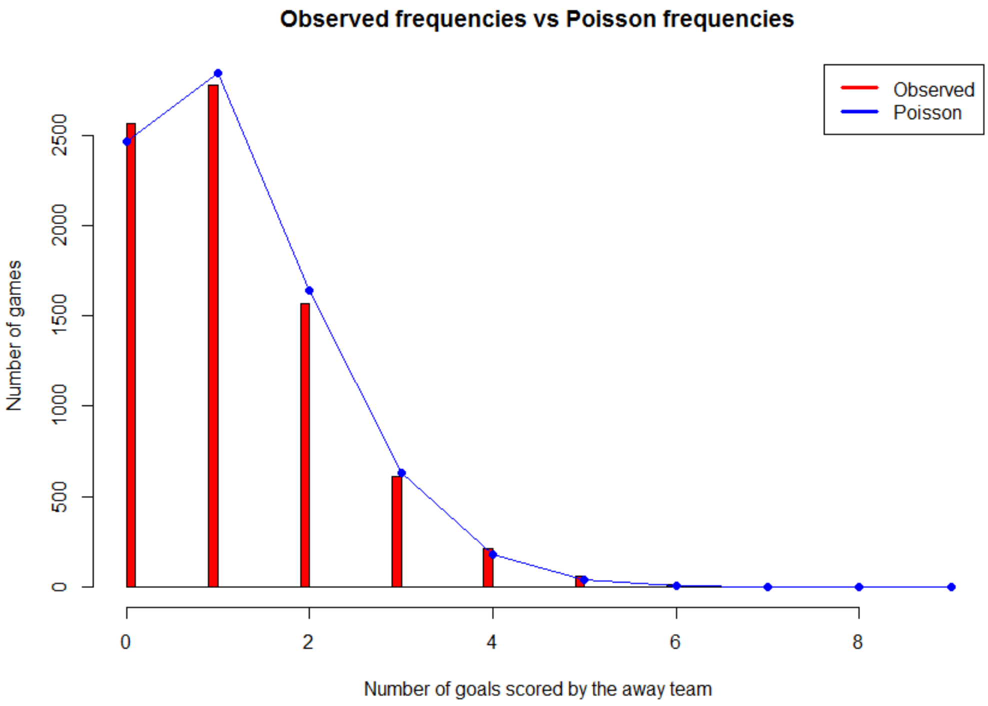


Figure 2: Plot comparing actual against expected Poisson frequencies for n.o. away goals

Away goals	Observed	Expected Poisson
0	2563	2464
1	2778	2843
2	1569	1639
3	609	630
4	214	182
5	59	42
6+	18	10

Table 4: Comparing observed against expected Poisson frequencies for n.o. awaygoals

By looking at Figure 1 and Figure 2, and also by comparing the observed with the expected Poisson frequencies in Table 3 and Table 4 it can be seen that the Poisson distribution fits the data quite well, although the observed distribution suggests a heavier tail than the Poisson. This is also a sign of overdispersion in the Poisson case, since the variance seems to be larger than the mean. The Poisson distribution assumption will be tested further using a goodness-of-fit test conducted in section 5.3.

5.2 Negative Binomial Distribution Assumption

An alternative to the Poisson distribution, which might prove to be a better fit for the data is the negative binomial distribution. Since the observed frequencies in Table 3 and Table 4 are more heavy tailed than the expected ones, there is some evidence that the Poisson distribution is overdispersed, i.e. that the variance is larger than the mean. Looking at Table 1, it can also be seen that the variance is indeed larger for most leagues. Assuming a Poisson distribution for a discrete variable is often too simplistic, since there are factors that easily cause overdispersion [3]. For example, the assumption that the probability of a goal being scored in a game is the same regardless of how long the game has progressed, as in the Poisson case, does not seem intuitively clear. It's more realistic that a goal being scored in a game is not entirely independent of time. A reasonable alternative to the Poisson distribution is therefore the negative binomial distribution, since it allows the variance to be larger than the mean. J. Greenhough et. al. [4] also suggests that football scores are approximated closer using a negative binomial distribution rather than the Poisson.

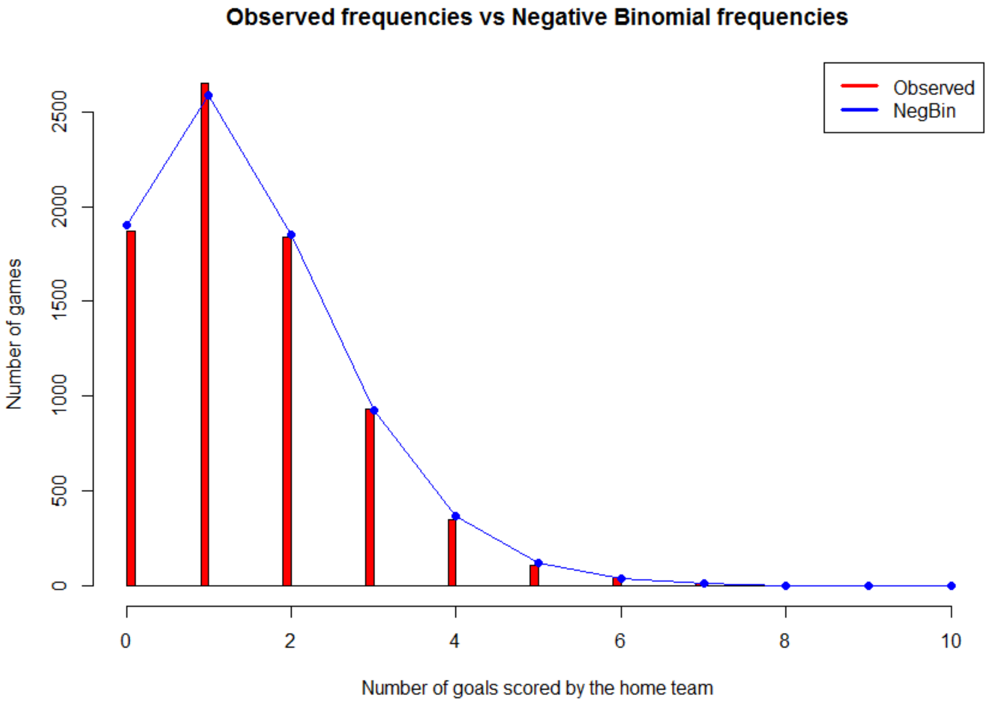


Figure 3: Plot comparing actual against expected Negative Binomial frequencies for n.o. home goals

Home goals	Observed	Expected negative binomial
0	1870	1905
1	2652	2591
2	1840	1854
3	933	929
4	349	365
5	107	120
6	45	34
7+	14	11

Table 5: Comparing observed against expected negative binomial frequencies for n.o. home goals

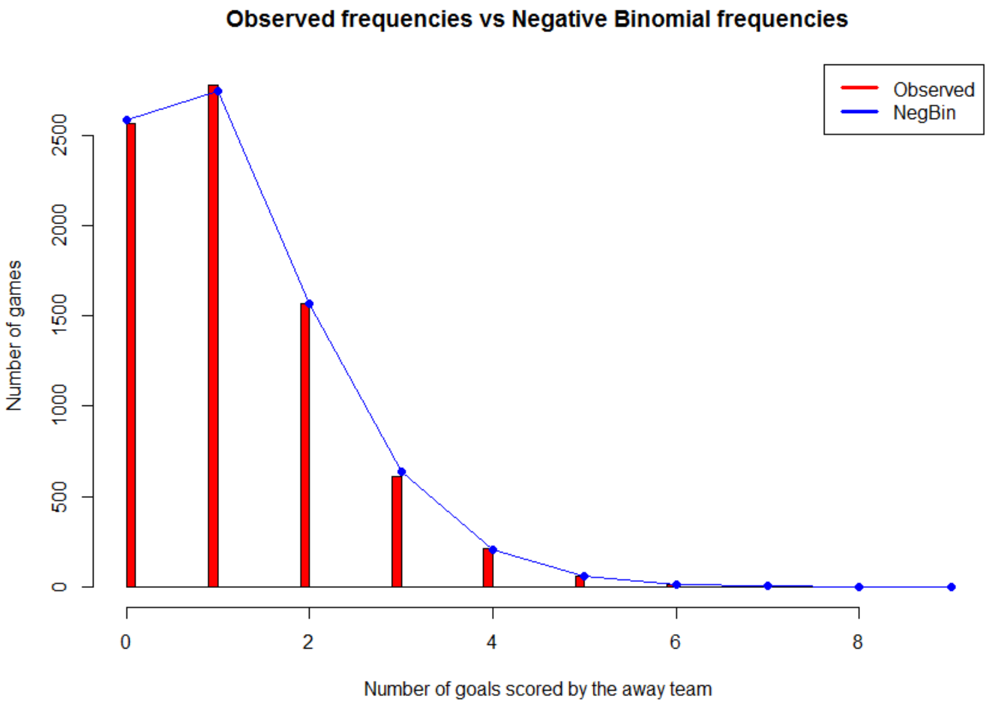


Figure 4: Plot comparing actual against expected negative binomial frequencies for n.o. away goals

Away goals	Observed	Expected negative binomial
0	2563	2583
1	2778	2742
2	1569	1565
3	609	637
4	214	207
5	59	57
6+	18	18

Table 6: Comparing observed against expected negative binomial frequencies for n.o. away goals

By looking at Figure 3 and Figure 4, it can be seen that the negative binomial distribution fits the data even better than the Poisson distribution. A goodness-of-fit test comparing both distributions to the data is conducted in section 5.3.

5.3 Goodness-of-fit Test

To investigate both the Poisson and the negative binomial distribution assumptions further, a chi-square goodness-of-fit test can be made. This test compares the observed frequencies, i.e. O_i to the expected theoretical frequencies given by a distribution, i.e. E_i. For this test, the test statistic has $c-p-1$ degrees of freedom, where c is the number of cathegories and p is the number of parameters used to estimate the distribution frequencies. The test statistic is computed as follows [3]:

$$X^2 = \sum_{i=1}^{c} \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{c-p-1}$$

In the Poisson case, one parameter is estimated, since λ is estimated to be the average of the number of goals made. In the negative binomial case however, two parameters are estimated. One parameter is μ which is estimated by the average number of goals scored, and the other one is a size parameter that is a target for number of successful trials. The goodness-of-fit tests were conducted using the R software and the code can be found in 8.1. All of the results can be seen in Table 7. All of the four tests have a significance level of 0.05.

Variable	Distribution assumption	p-value	Conclusion
N.o. home goals	Poisson	6.023828e-06	Reject
N.o. away goals	Poisson	3.541704e-05	Reject
N.o. home goals	Negative binomial	0.1221386	Accept
N.o. away goals	Negative binomial	0.7035373	Accept

Table 7: Goodness-of-fit tests comparing the Poisson- and negative binomial distribution to the observed data

By looking at Table 7, one can conclude that the Poisson assumption was rejected for both home and away goals, while the negative binomial assumption was accepted for both home and away goals. So the negative binomial distribution is indeed a better fit for the data than the Poisson distribution.

5.4 Independence Assumption

The two random variables representing home goals and away goals are assumed to be independent in the naive model. This is quite a strong assumption to make, and it's also possible to test whether this assumption holds true by using a Pearson's Chi-squared test. To test the hypothesis that the two random variables are indeed independent, data from 22 different European leagues have been used from the season 2015/2016. Pearson's Chi-squared test requires expected counts of at least five, which is why score counts of four or larger simply have been stated as "4+". Otherwise there would be frequencies in the table smaller than five. Investigating results from 7809 games during the football season 2015/2016, the following contingency table could be made:

Homegoals/Awaygoals	0	1	2	3	4+	Total
0	634	636	347	155	98	1870
1	842	954	547	212	97	2652
2	570	659	408	144	59	1840
3	318	333	181	70	31	933
4+	199	196	85	28	6	514
Total	2563	2778	1568	609	291	7809

Table 8: A table depicting the frequencies of games with a certain result during the 2015/2016 season

Now, using a Pearson's Chi-squared test in a contingency table with r rows and c columns, the test statistic is computed as follows [3]

$$X^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}} = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(n_{ij}-\hat{\mu}_{ij})^2}{\hat{\mu}_{ij}} \sim \chi^2_{(r-1)(c-1)}$$

$$\hat{\mu}_{ij} = n\hat{\pi}_{i+}\hat{\pi}_{+j} = \frac{n_{i+}n_{+j}}{n}$$

The test statistic is asymptotically chi-squared with $(r-1)(c-1)$ degrees of freedom. Using the R software, the null hypothesis that the two random variables home goals and away goals are independent can be tested. The p-value of the test comes out to be about 0.00004, so the null hypothesis is strongly rejected, since the p-value is a lot smaller than 0.05. Therefore, this test provides strong evidence that there is indeed an association between these two random variables. The R code used for the test can be found in section 8.1.

5.5 Fitting the Poisson Model

Since leagues differ in terms of aggressiveness and defensiveness, a certain number of matches will be used to predict the scoring intensities for home and away team. Teams change from season to season when players are bought and sold. A team manager might also be replaced by another one. Therefore the predicting data becomes less relevant the older it is. When running evaluations, the time interval for the selected data can also be varied to investigate what the optimal time frame for predictive data seems to be.

As an example, the number of homegoals in the Belgian Pro League during Season 2014/2015 have been modelled using the attack- and defense coefficients described in section 4.1 as predictors. In total, the model is based on 237 matches. An interaction term was dropped since it did not turn out to contribute significantly to explain the variation in the response variable. So using a Poisson Generalized Linear Model with the log-link, the following model is fitted:

$$\log(\lambda_{x,i}) = \alpha + \beta_1 attack_{x,i} + \beta_2 defense_{y,j}$$

$$\Rightarrow \lambda_{x,i} = \exp(\alpha + \beta_1 attack_{x,i} + \beta_2 defense_{y,j})$$

The summary output in R reports the following:

```
R output

Call:
glm(formula = prediction.data$FTHG ~ prediction.data$HomeAttack +
    prediction.data$AwayDefense, family = poisson(link = "log"))

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.5752  -0.7162  -0.1167   0.5481   2.5203

Coefficients:
                            Estimate Std. Error z value Pr(>|z|)
(Intercept)                  -1.4651     0.2636  -5.557 2.74e-08 ***
prediction.data$HomeAttack    0.9726     0.1949   4.990 6.05e-07 ***
prediction.data$AwayDefense   0.8601     0.1424   6.041 1.53e-09 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

```
(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 287.95  on 236  degrees of freedom
Residual deviance: 227.63  on 234  degrees of freedom
AIC: 696.49

Number of Fisher Scoring iterations: 5
```

The GLM calculates the parameters that maximize the likelihood. As can be seen in the R output, the intercept is highly significant as well as the coefficients for the home team's attack strength and the away team's defense strength. The coefficient for home attack is positive, which is also intuitive, since a home team with a large attack coefficient (i.e. a strong team) must be expected to score more goals. It can also be seen that the coefficient for away defense is positive, which also makes intuitive sense, because a weaker team has a large defense coefficient since it concedes many goals. So an increase in the away team's defense coefficient must yield an increase in the expected number of goals scored by the home team. To determine if the model shows evidence of overdispersion, the ratio of the residual deviance over the degrees of freedom can be evaluated. The R output yields:

$$\frac{\text{Residual deviance}}{\text{Degrees of freedom}} = \frac{227.63}{234} = 0.973 < 1$$

So there is no direct evidence of overdispersion in the model. To test whether the model is a good fit for the data, a deviance goodness-of-fit test can be made, explained in section 4.4. Using the R software to test the null hypothesis that the model is an appropriate fit for the data yields a p-value of 0.61, so the null hypothesis that the model is not an appropriate fit for the data, is not rejected. The residual deviance is still quite large, so there is still a lot of variability in the response variable that the home attack- and away defense coefficients cannot explain. The R code used to conclude this test can be found in section 8.1.

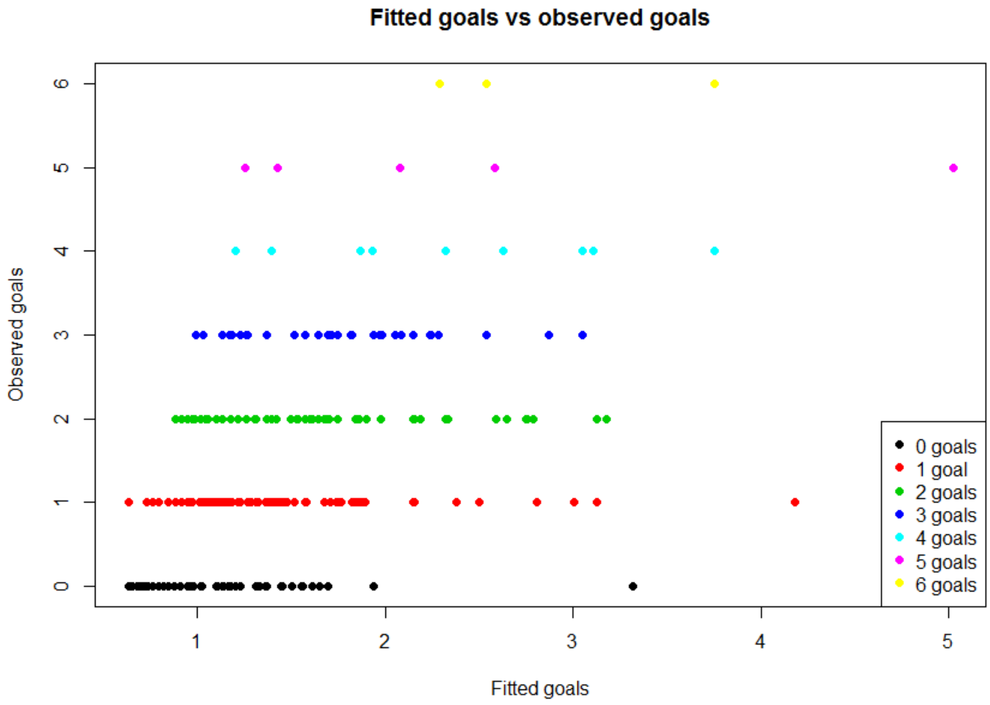


Figure 5: Fitted goals plotted against observed goals for the home team

Figure 5 shows the fitted expected number of goals for the home team against the observed number of goals for the home team. Since the observed data is discrete which takes on integer values from 0 to 6, the points are aligned along different numbers of observed goals. A pattern can be spotted where the observed number of goals seem to increase with the fitted number of goals, i.e. a positive correlation. This is expected if the model has any predictive power at all. The fitted goals are, however, quite spread out for each alignment of observed goals.

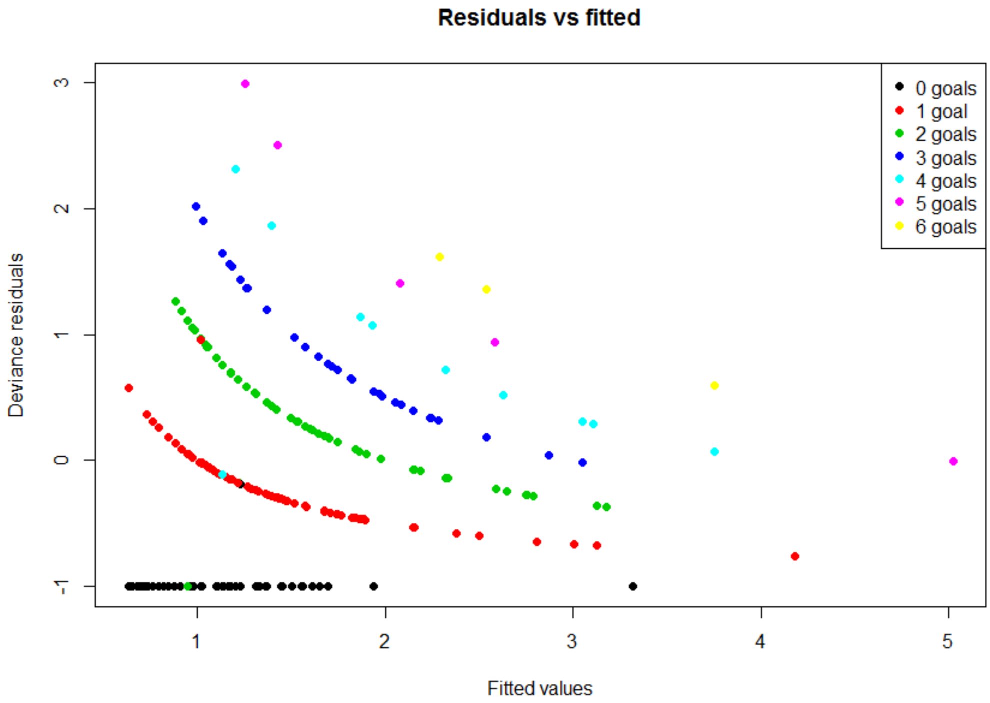


Figure 6: Fitted goals plotted against deviance residuals for the home team

In the residual plot in Figure 6, clear patterns can be spotted as the deviance residuals seem to decrease exponentially with the fitted values. This is to expect since the log link is being used. The color coded points represent the number of goals for each observation. No home team scored more than six goals in any match. Another important aspect of assessing a model is whether multicollinearity seems to be present or not. If the model exhibits multicollinearity, the two predictors are strongly correlated to one another. In this case, dropping one of them can lead to more precise estimates of the parameters, since the standard error decreases [3]. To do so, the two predictors have been plotted against each other. The result can be viewed in Figure 7.

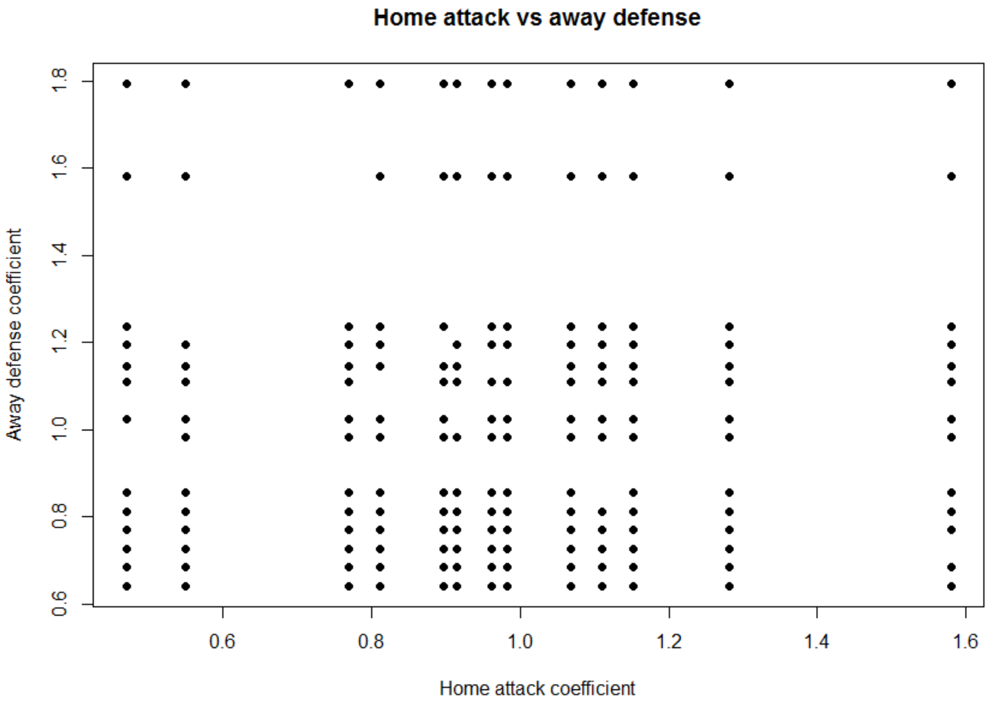


Figure 7: Home attack coeffienct plotted against away defense coefficient

As can be seen in Figure 7, the data points align themselves since many teams have the exact same home attack or away defense coefficient, since the teams in question have scored or conceded the same number of goals. There does not seem to be a linear relationship between the two predictors, so there is no evidence of multicollinearity. Also, the correlation between the two predictors is calculated as 0.025, which is close to zero.

Now, for example, the home team Standard Liege is going to play against the away team Charleroi. Using the fitted model, the expected number of goals for Standard Liege can be calculated. In this case, Standard Liege has a home attack coefficient of 1.11 and Charleroi has an away defense coefficient equal to 0.81. Therefore, the expected number of goals for Standard Liege according to the model is:

$$\lambda_{x,Standard} = \exp(\alpha + \beta_1 attack_{x,Standard} + \beta_2 defense_{y,Charleroi})$$

$$\Rightarrow \lambda_{x,Standard} = \exp(-1.4651 + 0.9726 * 1.11 + 0.8601 * 0.81) = 1.37$$

So the home team Standard Liege is expected to score 1.37 goals against the away team Charleroi. The expected goals for the away team is modelled in a completely analogous way, but instead uses the home team's defense coefficient and the away team's attack coefficient as predictors. Using the same 237 matches as prediction data, fitting a GLM for the number of away goals yields the following model:

$$\log(\lambda_{y,j}) = \alpha + \beta_1 attack_{y,j} + \beta_2 defense_{x,i}$$

$$\Rightarrow \lambda_{y,j} = \exp(\alpha + \beta_1 attack_{y,j} + \beta_2 defense_{x,i})$$

Fitting a GLM in R gives the following output:

```
Call:
glm(formula = prediction.data$FTAG ~ prediction.data$AwayAttack +
    prediction.data$HomeDefense, family = poisson(link = "log"))

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.2169  -1.1866  -0.1300   0.5947   2.8146

Coefficients:
                            Estimate Std. Error z value Pr(>|z|)
(Intercept)                  -1.7863     0.2799  -6.382 1.75e-10 ***
prediction.data$AwayAttack    0.9190     0.1662   5.528 3.24e-08 ***
prediction.data$HomeDefense   0.9666     0.1883   5.133 2.85e-07 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 315.65  on 236  degrees of freedom
Residual deviance: 257.45  on 234  degrees of freedom
AIC: 644.39

Number of Fisher Scoring iterations: 5
```

As can be seen, all estimated parameters including the intercept are highly significant. The parameter representing the attack coefficient of the away team is positive, which is intuitive since an increase in the away team's attack coefficient must yield an increase in the expected number of goals for the away team, since the away team in that case scores more goals. In a similar fashion, an increase in the home team's defense coefficient means that the home team is weaker, since it concedes more goals. Therefore it's also intuitive that the parameter representing the defense coefficient of the home team is also positive. Checking for overdispersion described in section 4.5, yields the following:

$$\frac{\text{Residual deviance}}{\text{Degrees of freedom}} = \frac{257.45}{234} = 1.10 > 1$$

So in this case, there is some evidence of overdispersion, so the Poisson assumption is not quite adequate, and a negative binomial GLM is preferable. To further assess the validity of the model, a deviance goodness-of-fit test can be made, which is described in section 4.4. Testing the null hypothesis that the model is an adequate fit for the data yields a p-value of about 0.14, which is not small enough to reject the null, but a larger p-value would be preferable. The residual deviance is still quite large, just as in the case of the model of the home goals, so a smaller residual deviance with significant predictors would be preferable. The R code used can be found in section 8.1. A plot of the fitted goals against the observed goals can be found in Figure 8.

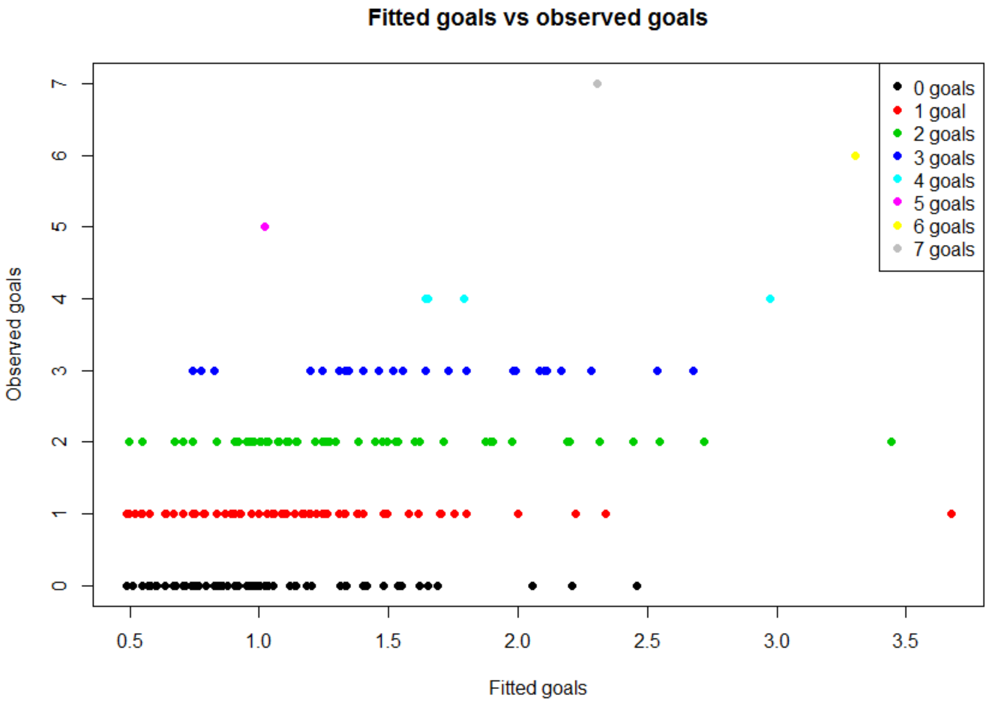


Figure 8: Fitted goals plotted against observed goals for the away team

As can be seen in Figure 8, the observed number of goals take on integer values from 0 to 7, creating an alignment along the number of observed goals. There seems to be a positive correlation between fitted goals and observed goals, which is expected if the model has any predictive power at all. In that case, the number of observed goals are expected to increase if the number of fitted goals increases. R reports a correlation of about 0.47, which is quite low, since there is a lot of variability along each alignment of observed goals. A perfect model that predicted the correct number of goals for each fitted value would of course yield a correlation of 1.

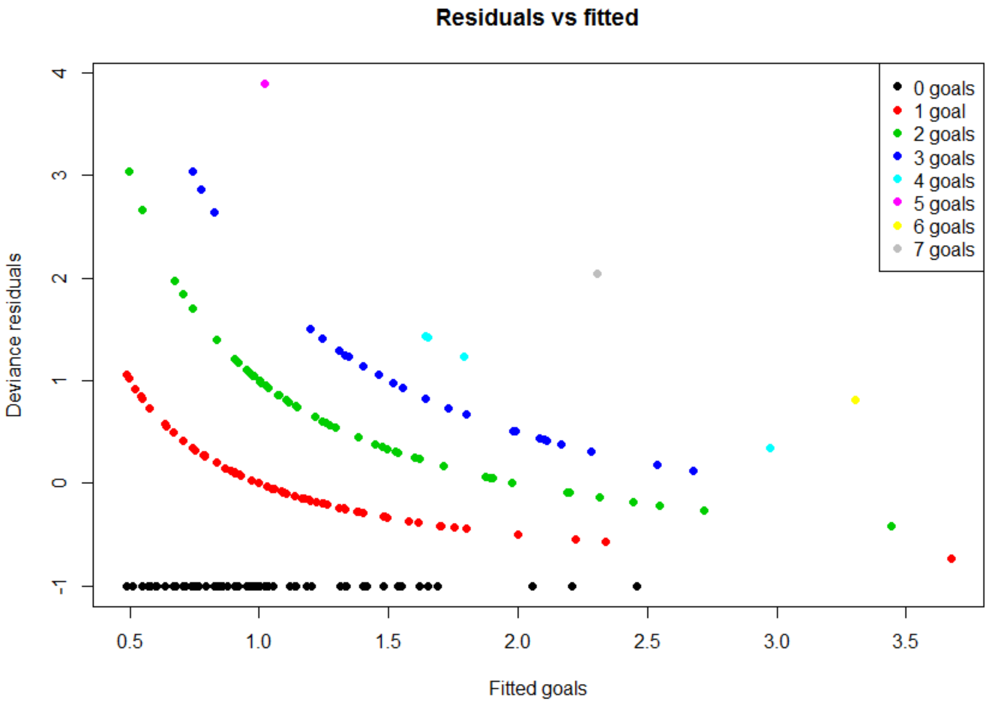


Figure 9: Fitted goals plotted against deviance residuals for the away team

In Figure 9 the fitted away goals are plotted against the deviance residuals. A similar pattern as in the case for the home goals can be spotted, since the deviance residuals seem to decrease exponentially as the number of fitted goals increase, which is intuitive since the log link is used. Finally, it's also important to check the model for multicollinearity, by plotting the away attack coefficient against the home defense coefficient.

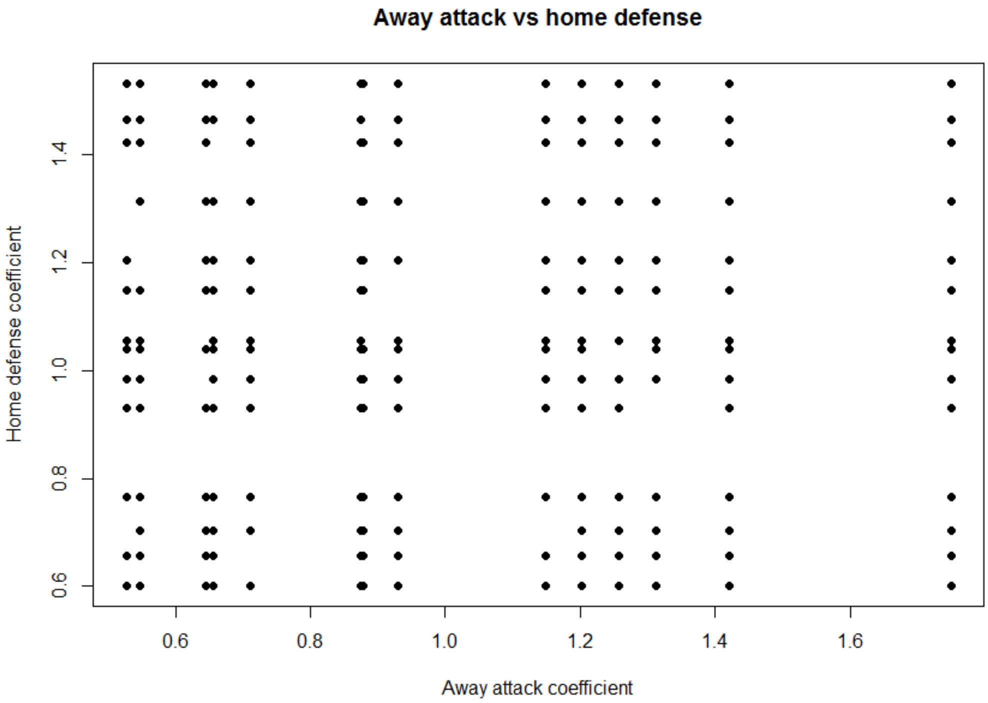


Figure 10: Fitted goals plotted against deviance residuals for the away team

Figure 10 does not show any apparent evidence of multicollinearity, since there does not seem to be a noticable correlation between the away team's attack coefficient and the home team's defense coefficient. The R software reports a correlation of about 0.03, which is close to zero. As can be seen, the discrete points align themselves since multiple teams have the exact same away attack coefficient or home defense coefficient since they score or concede the same number of goals.

Now, in the match between Standard Liege and Charleroi, the expected number of away goals (i.e. the expected goals for Charleroi) can be computed using the model. Since Standard Liege has a home defense coefficient of 1.04 and Charleroi has an away attack coefficient of 1.15, the expected number of goals that Charleroi scores can be computed as follows:

$$\lambda_{y,Charleroi} = \exp(\alpha + \beta_1 attack_{y,Charleroi} + \beta_2 defense_{x,Standard})$$

$$\Rightarrow \lambda_{y,Charleroi} = \exp(-1.7863 + 0.9190 * 1.15 + 0.9666 * 1.04) = 1.32$$

So Charleroi is expected to score 1.32 goals away against Standard Liege. As seen earlier, the expected number of goals for the home team Standard Liege came out to be 1.37. Since both expectations are computed the joint probability mass function can be used to compute probabilities of all results, and therefore estimate probabilities for a home win, draw and an away win.

5.6 Fitting the Negative Binomial Model

Since the goodness-of-fit test in section 5.3 showed that the negative binomial distribution is in fact a better fit for the data than the Poisson distribution, a negative binomial GLM is also fitted for the same 237 matches in the Belgian Pro League during season 2014/2015. The GLM uses the home team's attack coefficient and the away team's defense coefficient as predictors just as in the Poisson model. The R output from the fitted model using number of home goals as the response variable can be read as follows:

```
Call:
glm.nb(formula = prediction.data$FTHG ~ prediction.data$HomeAttack +
    prediction.data$AwayDefense, link = log, init.theta = 23005.47029)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.5752  -0.7161  -0.1167   0.5481   2.5202

Coefficients:
                            Estimate Std. Error z value Pr(>|z|)
(Intercept)                  -1.4652     0.2637  -5.557 2.74e-08 ***
prediction.data$HomeAttack    0.9726     0.1949   4.990 6.05e-07 ***
prediction.data$AwayDefense   0.8601     0.1424   6.041 1.53e-09 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for Negative Binomial(23005.47) family taken to be 1)

    Null deviance: 287.93  on 236  degrees of freedom
Residual deviance: 227.61  on 234  degrees of freedom
AIC: 698.49

Number of Fisher Scoring iterations: 1


              Theta:  23005
          Std. Err.:  239354
Warning while fitting theta: iteration limit reached

 2 x log-likelihood:  -690.495
```

As can be seen, the parameter estimates are almost exactly the same as in the Poisson model. Another interesting thing to notice is that the theta parameter seems to diverge ($\theta = k = 23005$) in the Fisher Scoring Algorithm, which is an iterative method of solving maximum likelihood equations [3]. Since $\theta = k$, i.e. the specified number of failures, the negative binomial converges to the Poisson distribution as discussed in section 4.5.

This is expected since the model does not show any sign of overdispersion, since the following ratio can be observed:

$$\frac{\text{Residual deviance}}{\text{Degrees of freedom}} = \frac{227.61}{234} < 1$$

Since there is no apparent evidence of overdispersion, the Poisson assumption is adequate for this particular data. However, using different data might cause overdispersion, so using a negative binomial GLM is more general overall. The Poisson GLM and the negative binomial GLM have almost exactly the same parameter estimates, which means that the estimate of the expected number of home goals will not change much whether Poisson regression or negative binomial regression is used. This also means that the residual plot and the ”Fitted vs observed” plot both look almost exactly identical as in the Poisson case. A deviance goodness-of-fit test, described in section 4.4, can be made for the negative binomial model as well, yielding a p-value of 0.61, so the null hypothesis that the model is a good fit for the data cannot be rejected, even though the residual deviance is still quite large, meaning that there is still a lot of variability in the response variable that cannot be explained by the home attack- and away defense coefficient. The R code used can be found in section 8.1.

To model the away goals using the away team's attack coefficient and the home team's defense coefficient, the following negative binomial GLM was fitted in R:

```
Call:
glm.nb(formula = prediction.data$FTAG ~ prediction.data$AwayAttack +
    prediction.data$HomeDefense, link = log, init.theta = 12225.57363)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.2168  -1.1866  -0.1300   0.5946   2.8144

Coefficients:
                            Estimate Std. Error z value Pr(>|z|)
(Intercept)                  -1.7863     0.2799  -6.381 1.75e-10 ***
prediction.data$AwayAttack    0.9190     0.1662   5.528 3.24e-08 ***
prediction.data$HomeDefense   0.9666     0.1883   5.133 2.85e-07 ***
```

```
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for Negative Binomial(12225.57) family taken to be 1)

    Null deviance: 315.62  on 236  degrees of freedom
Residual deviance: 257.43  on 234  degrees of freedom
AIC: 646.39

Number of Fisher Scoring iterations: 1


              Theta:  12226
          Std. Err.:  155027
Warning while fitting theta: iteration limit reached

 2 x log-likelihood:  -638.391
```

As can be seen, the parameter estimates are almost exactly the same as in the Poisson GLM, meaning that the fitted vs observed plot as well as the residual plot will look almost exactly the same. In the case of overdispersion, the following ratio can be observed:

$$\frac{\text{Residual deviance}}{\text{Degrees of freedom}} = \frac{257.43}{234} = 1.10 > 1$$

So there is some evidence of overdispersion, but not much. It can also be seen that $\theta = k = 12226$, which indicates very little overdispersion as discussed earlier. A deviance goodness-of-fit test reports a p-value of 0.14, so it cannot be rejected that the model is a good fit for the data. However, the residual deviance is still quite large, so there is a lot of variability in the response variable that cannot be explained by the away attack coefficient and the home defense coefficient. To summarize, there is almost no difference between the two Poisson models and the two negative binomial models. They both yield approximately the exact same parameter estimates, however the models are fitted to a very small subset of all data, so just because the two different GLM models are very similar doesn't necessarily mean that it will be the case for all models fitted, since different data will yield different parameter estimates.

5.7 Results of Betting Evaluations

To conclude which one of the models that performs best in the long run, and to also investigate whether any one of them is accurate enough to profit in the long run, bets have been made using the evaluation program described in section 4.8. Using an error margin of $r = 0.05$, and a time range of 365 days, so that the evaluation program selects prediction data from the same league during the last year, yields an evaluation depicted in Figure 11. Here, an initial bankroll of 1000 fictional units is assumed, and for all models a bet size of $\frac{20}{\text{Maximum odds}}$ is used, since bets with high odds have larger variance and bets with low odds have smaller variance. Using a bet size relative to the odds is therefore implemented in order to decrease the probability of ruin. Random bets are also made, where the evaluation program simply randomizes which of the three outcomes to bet on with probabilities equal to the relative frequencies of the three outcomes during the season, and then selects the maximum odds to bet on. Relative frequencies of home wins, draws and away wins can be found in Table 9.

Frequencies/Outcomes	Home wins	Draws	Away wins	Totals
Frequencies	3412	2071	2327	7810
Relative frequencies	0.437	0.265	0.298	1

Table 9: Absolute and relative frequencies during the 2015/2016 season

As can be seen in Table 9, the relative frequencies for home wins, draws and away wins are 0.437, 0.265 and 0.298. These relative frequencies are used to estimate the respective probability for each of the three outcomes that the evaluation program then uses to select a random outcome to bet on.

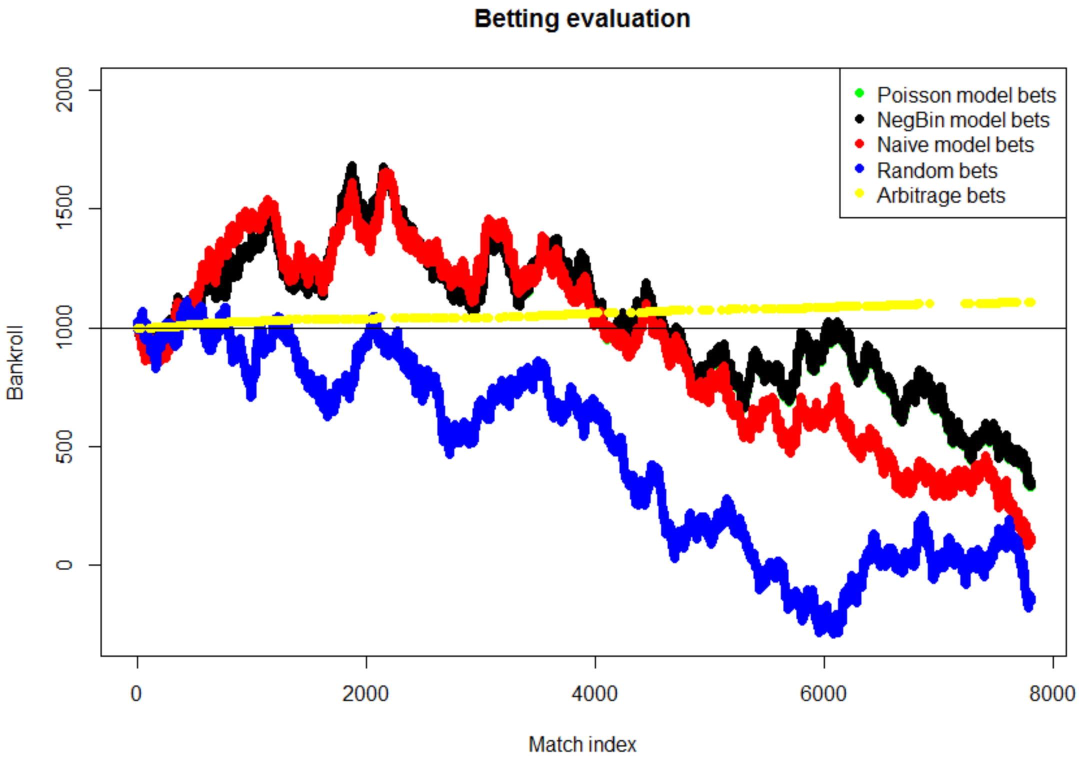


Figure 11: Bankroll fluctuations of different models and random bets during the 2015/2016 season (r=0.05, time frame=365 days)

As can be seen in Figure 11, the bets using the Poisson model are exactly the same as the bets using the negative binomial model. In section 5.5 and 5.6, it can be seen that the parameters in both the fitted Poisson and the negative binomial GLM are almost exactly equal to each other, yielding almost identical probability matrices and therefore the two models will make almost the exact same bets. So therefore the bankroll fluctuation of the Poisson model is not visible, since the two models overlap. The naive model follows the Poisson and negative binomial models closely, which is intuitive since it also uses a bivariate Poisson distribution when computing probabilities, although it assumes independence between home goals and away goals, as described in section 4.1. All three non random models yield a loss after about 7000 bets, so none of the models seem to predict probabilities accurately enough to profit in the long run. The random bets cause bankruptcy, which is also intuitive, since simply betting randomly is not expected to be a profitable strategy. The only profitable strategy judging by Figure 11 seems to be the arbitrage strategy. In the figure, it is also visible that the arbitrage bets have ”jumps” in between bets, since bets are not made for each match in the 2015/2016 season, but only in games where arbitrage opportunities exist. The prediction models also have conditions that need to be met in order for a bet to be placed, which also prevent the models from betting on each match.

Model	**Iterations**	**N.o. bets**	**Profit**	**Total amount bet**	**Profitmargin**	**r**	**Timeframe (days)**	**St.Dev. Profit**
Poisson	7800	7003	-669	45909	-0.0146	0.05	365	8.8929
NegBin	7800	7002	-665	45905	-0.0145	0.05	365	8.8934
Naive	7800	7025	-902	45905	-0.0196	0.05	365	9.0080
Random	7800	7634	-1139	54359	-0.0210	0.05	365	8.9853
Arbitrage	7800	**666**	**108**	**13103**	**0.0082**	**0.05**	**365**	**0.1505**
Poisson	7800	6109	-442	38882	-0.0114	0.10	365	8.8648
NegBin	7800	6109	-442	38882	-0.0114	0.10	365	8.8648
Naive	7800	6161	-642	41502	-0.0155	0.10	365	8.9889
Random	7800	7634	-2166	54126	-0.040	0.10	365	9.0137
Arbitrage	**7800**	**666**	**108**	**13103**	**0.0082**	**0.10**	**365**	**0.1505**
Poisson	7800	5180	-455	32055	-0.0142	0.15	365	8.8171
NegBin	7800	5179	-462	32042	-0.0144	0.15	365	8.8175
Naive	7800	5266	-212	34352	-0.0062	0.15	365	8.9439
Random	**7800**	**7634**	**718**	**54222**	**0.0132**	**0.15**	**365**	**9.0921**
Arbitrage	**7800**	**666**	**108**	**13103**	**0.0082**	**0.15**	**365**	**0.1505**
Poisson	7800	4330	-365	25885	-0.0141	0.20	365	8.7475
NegBin	7800	4330	-365	25885	-0.0141	0.20	365	8.7475
Naive	7800	4475	-258	28278	-0.009	0.20	365	8.8659
Random	**7800**	**7634**	**406**	**54434**	**0.0075**	**0.20**	**365**	**9.1272**
Arbitrage	**7800**	**666**	**108**	**13103**	**0.0082**	**0.20**	**365**	**0.1505**
Poisson	7800	3615	-440	20940	-0.0210	0.25	365	8.6715
NegBin	7800	3616	-450	20950	-0.0215	0.25	365	8.6717
Naive	7800	3865	-162	23682	-0.007	0.25	365	8.8280
Random	7800	7634	-853	53953	-0.0158	0.25	365	9.0014
Arbitrage	**7800**	**666**	**108**	**13103**	**0.0082**	**0.25**	**365**	**0.1505**

Table 10: A table depicting the outcomes of evaluations of different models during the 2015/2016 season, varying the error margin against the bookmakers (r) with a time frame of 365 days

Table 10 shows key values of outcomes of different evaluations. Profitable results have been indicated with a bold font. The number of iterations is simply how many games that have been predicted using the different models. Since the evaluation program does not bet if the attack or the defense coefficient is equal to zero or if r is not larger than the constant given in Table 10, bets are not made for every single game. In the evaluations carried out, the error margin of the calculated odds compared to the actual odds (r) has been varied using a time frame of 365 days. So each prediction of the outcome of a match is determined by 365 days of historical results prior to the match being played. None of the predictive models manage to make a profit regardless of the value of r, so increasing the error margin does not seem to improve the results, although the smallest relative loss for the Poisson and negative binomial models occurs when $r = 0.10$. As can be seen, the only consistent winner is the arbitrage strategy, which always makes the same bets in every evaluation, since the data is fixed and the strategy does not depend on the varied parameters. Even the random bets manage to make a profit in two of the evaluations, which is due to pure variance, as simply betting randomly cannot possibly be a profitable strategy in the long run.

Model	Iterations	N.o. bets	Profit	Total amount bet	Profitmargin	r	Timeframe (days)	St.Dev. Profit
Poisson	**7800**	**6530**	**161**	**43519**	**0.0037**	**0.10**	**73**	**9.0569**
NegBin	**7800**	**6530**	**161**	**43519**	**0.0037**	**0.10**	**73**	**9.0569**
Naive	7800	6825	-562	48482	-0.0116	0.10	73	9.0963
Random	**7800**	**7351**	**126**	**51854**	**0.0024**	**0.10**	**73**	**9.1144**
Arbitrage	**7800**	**634**	**101**	**12476**	**0.0081**	**0.10**	**73**	**0.1496**
Poisson	7800	6563	-297	43777	-0.0068	0.10	146	8.9755
NegBin	7800	6563	-297	43777	-0.0068	0.10	146	8.9755
Naive	7800	6751	-817	47677	-0.0171	0.10	146	9.0532
Random	7800	7597	-915	53575	-0.0171	0.10	146	9.0446
Arbitrage	**7800**	**663**	**108**	**13043**	**0.0082**	**0.10**	**146**	**0.1506**
Poisson	7800	6307	-801	41681	-0.0192	0.10	219	8.9191
NegBin	7800	6307	-801	41681	-0.0192	0.10	219	8.9191
Naive	7800	6499	-826	45246	-0.0183	0.10	219	9.0284
Random	7800	7633	-1620	54120	-0.0299	0.10	219	8.9950
Arbitrage	**7800**	**666**	**108**	**13103**	**0.0082**	**0.10**	**219**	**0.1505**
Poisson	7800	6210	-507	40327	-0.0126	0.10	292	8.8727
NegBin	7800	6210	-507	40327	-0.0126	0.10	292	8.8727
Naive	7800	6329	-550	43350	-0.0127	0.10	292	9.0040
Random	7800	7634	-86	53846	-0.0016	0.10	292	9.0257
Arbitrage	**7800**	**666**	**108**	**13103**	**0.0082**	**0.10**	**292**	**0.1505**
Poisson	7800	6109	-442	38882	-0.0114	0.10	365	8.8648
NegBin	7800	6109	-442	38882	-0.0114	0.10	365	8.8648
Naive	7800	6161	-642	41502	-0.0155	0.10	365	8.9889
Random	7800	7634	-725	54005	-0.0134	0.10	365	9.0604
Arbitrage	**7800**	**666**	**108**	**13103**	**0.0082**	**0.10**	**365**	**0.1505**
Poisson	7800	6040	-451	38451	-0.0117	0.10	438	8.8667
NegBin	7800	6041	-454	38454	-0.0118	0.10	438	8.8661
Naive	7800	6082	-244	40324	-0.0061	0.10	438	8.9442
Random	7800	7634	-1139	54359	-0.0209	0.10	438	8.9853
Arbitrage	**7800**	**666**	**108**	**13103**	**0.0082**	**0.10**	**438**	**0.1505**
Poisson	7800	5984	-366	38026	-0.0096	0.10	511	8.8542
NegBin	7800	5983	-398	38018	-0.0105	0.10	511	8.8509
Naive	7800	6042	-1005	39905	-0.0252	0.10	511	8.9157
Random	7800	7673	-659	54439	-0.0121	0.10	511	9.0420
Arbitrage	**7800**	**668**	**108**	**13142**	**0.0082**	**0.10**	**511**	**0.1503**
Poisson	7800	5962	-551	37651	-0.0146	0.10	584	8.8394
NegBin	7800	5961	-565	37645	-0.0150	0.10	584	8.8381
Naive	7800	6050	-840	39600	-0.0212	0.10	584	8.9026
Random	7800	7674	-1089	54589	-0.0200	0.10	584	9.0407
Arbitrage	**7800**	**668**	**108**	**13142**	**0.0082**	**0.10**	**584**	**0.1503**
Poisson	7800	5897	-702	37162	-0.0189	0.10	657	8.8399168
NegBin	7800	5895	-691	37151	-0.0186	0.10	657	8.8409
Naive	7800	5974	-999	39139	-0.0255	0.10	657	8.8953
Random	7800	7674	-824	54264	-0.0152	0.10	657	9.0661
Arbitrage	**7800**	**668**	**108**	**13142**	**0.0082**	**0.10**	**657**	**0.1503**

Table 11: A table depicting the outcomes of evaluations of different models during the 2015/2016 season, using an the error margin of 0.10 (r=0.10) and a varying time frame

In Table 11, results of more betting evaluations can be seen, varying the time frame while using a consistent error margin of 0.10. As can be seen, the Poisson and negative binomial models manage to make a small profit using a small time frame of 73 days, but otherwise none of the prediction models manage to make profit for any other time frame. As mentioned, the arbitrage strategy makes the same bets regardless of the parameters, and therefore makes the same profit in every evaluation. Here, even longer time frames up to 657 days have been used, although increasing the time frame does not seem to improve the results of the evaluations. Actually, the best results of the Poisson and negative binomial models occur for a small time frame of 73 and 146 days. However, there is still a lot of variability in the results of the evaluations, which the random bets indicate, so conclusions must be interpreted carefully. The respective plots of the betting evaluations in both Table 10 and Table 11 can be found in section 8.2.

6 Discussion

From the results of the betting evaluations described in section 5.7, it can be concluded that since almost none of the evaluations using the prediction models (i.e. the Poisson, negative binomial and the naive model) managed to make a profit in the long run, these models do not seem to predict probabilities of results more accurately than the bookmakers. Of course, the models are also quite simple, depending only on attack- and defense coefficients (and in the Poisson and negative binomial case also a dependence parameter) for home- and away team. The models do not take into account other variables that might affect the probability of a certain result, for example how a team has been performing recently, i.e. if a team is in good form or not. Including also such a variable might have a significant impact on the expectation of number of home- and away goals. There are of course also many other variables that the models do not take into account such as injuries, weather conditions, travel distance for the away team etc. Also, if more predictors are to be included in the GLM's, they must also contribute significantly to explain the variability in the response variable. Including many non-significant predictors would cause the standard errors of the fitted regression parameters to increase, and might do more harm than good. All in all, it seems like the only strategy which actually manages to make a long term profit is the arbitrage strategy, even though the profit margin is very small. However, since arbitrage bets theoretically do not include any risk, it is possible to bet a lot more on each arbitrage opportunity, which would yield a larger profit in absolute terms. In the real world, however, arbitrage bets still include some degree of risk, since bookmakers might reject or cancel a bet made, so if that outcome is not covered at another betting site, then there is a risk of losing the bets on the other two outcomes.

Moreover, assuming a negative binomial distribution for the number of home- and away goals instead of a Poisson distribution does not make much difference. This is also intuitive since the overdispersion in the data is quite small. Also, choosing fewer cathegories for the number of home- and away goals in the goodness-of-fit tests might have led to accepting the Poisson distribution assumption. From the betting evaluations, it can also be concluded that the Poisson model (which uses GLM's to estimate expectations and also assumes dependence between home- and away goals) did not perform much better than the naive model which also assumes Poisson distribution but independence between home- and away goals. The naive model also differs from the Poisson model since it does not use GLM's to estimate expectations. The overall dependence between home- and away goals is quite weak, with a small negative correlation. Since the dependence is weak, a model assuming dependence instead of independence might not improve the accuracy of the predictions to a large extent. Using GLM's to estimate expectations instead of simply using the attack- and defense coefficients as is the case in the naive model described in section 4.1, does not seem to improve the betting results much. The residual deviances in the GLM's are still quite large, meaning that there is still a lot of variability in the home- and away goals which cannot be explained simply by the attack- and defense coefficients.

To see if results differ significantly or if the loss made by the prediction models is significant, it would be interesting to use hypothesis testing. In that case a hypothesis test could be made where the null hypothesis is a zero profit for a certain model and the alternative hypothesis is that the profit is less than zero. However, the bets have not been selected randomly and therefore the conditions of a hypothesis test are not met, so it can simply not be performed. Another interesting test would be to test if the prediction models can significantly outperform the random bets (i.e. if the loss of the bets made by the prediction models is significantly smaller than the loss of the random bets). Unfortunately, the problem with bets not being randomly selected persists. The bets have also not been selected independently of one another, which is another condition for a hypothesis test. Since there is a time dependence between the games, a method that could be used when working with dependent data (as in this case) is to use non-overlapping block bootstrap. In this case the data is split into non-overlapping blocks of a specific length. Then, by sampling the blocks with replacement, assuming that the blocks are independent of one another, since games played far apart from one another can be assumed to be approximately independent, a distribution of the test statistic can be found and therefore a hypothesis test can be concluded. However, such a procedure would not be very time efficient, since it would require many betting evaluations made by the program.

7 References

References

[1] Pinnacle (2014). https://www.pinnacle.com/en/betting-articles/soccer/how-to-calculate-poisson-distribution. 2016-09-07.

[2] http://www.football-data.co.uk/downloadm.php. 2016-09-09.

[3] Agresti, Alan. (2013). Categorical Data Analysis. Hoboken, New Jersey: John Wiley Sons.

[4] J. Greenhough, P.C. Birch, S.C. Chapman, G. Rowlands (2002). Football goal distributions and extremal statistics. Department of Physics, University of Warwick. https://www2.warwick.ac.uk/fac/sci/physics/research/cfsa/people/sandrac/publications/footy.pdf. 2016-09-12.

[5] Cameron, A. Colin, Trivedi, Pravin K. (2013). Regression Analysis of Count Data. New York, NY: Cambridge University Press.

[6] Albert, Jim. Koning, Ruud H. (2007). Statistical Thinking in Sports. Boca Raton, FL: Taylor Francis Group, LLC.

[7] van Gemert, Dan. (2010). Modelling the scores of Premier League Football Matches. Department of Quantitative Economics, University of Amsterdam. http://dare.uva.nl/cgi/arno/show.cgi?fid=175669. 2016-09-29.

[8] S. N. Lahiri. (2003). Resampling Methods for Dependent Data. New York: Springer-Verlag.

8 Appendix

8.1 R Code

```
> #Goodness-of-fit test Poisson home goals
>
> poisson.home.freqs<-c(1806, 2645, 1936, 945, 346, 101, 25, 6);
> observed.freqs<-c(1870, 2652, 1840, 933, 349, 107, 45, 14);
>
> chisq.test.statistic<-sum((observed.freqs-poisson.home.freqs)^2
/poisson.home.freqs);
>
> df<-length(observed.freqs)-2;
>
> p.value<-pchisq(chisq.test.statistic, df, lower.tail = FALSE);
> p.value
[1] 6.023828e-06
```

```
> #Goodness-of-fit test Poisson number of away goals
>
> poisson.away.freqs<-c(2464, 2843, 1639, 630, 182, 42, 10);
> observed.freqs<-c(2563, 2778, 1569, 609, 214, 59, 18);
>
> chisq.test.statistic<-sum((observed.freqs-poisson.away.freqs)^2
/poisson.away.freqs);
>
> df<-length(observed.freqs)-2;
>
> p.value<-pchisq(chisq.test.statistic, df, lower.tail = FALSE);
> p.value
[1] 3.541704e-05
```

```
> #Goodness-of-fit test NegBin number of home goals
>
> negbin.home.freqs<-c(1905, 2591, 1854, 929, 365, 120, 34, 11);
> observed.freqs<-c(1870, 2652, 1840, 933, 349, 107, 45, 14);
>
> chisq.test.statistic<-sum((observed.freqs-negbin.home.freqs)^2
/negbin.home.freqs);
>
> df<-length(observed.freqs)-3;
```

```
>
> p.value<-pchisq(chisq.test.statistic, df, lower.tail = FALSE);
> p.value
[1] 0.1221386
```

```
> #Goodness-of-fit test NegBin number of away goals
>
> negbin.away.freqs<-c(2583, 2742, 1565, 637, 207, 57, 18);
> observed.freqs<-c(2563, 2778, 1569, 609, 214, 59, 18);
>
> chisq.test.statistic<-sum((observed.freqs-negbin.away.freqs)^2
/negbin.away.freqs);
>
> df<-length(observed.freqs)-3;
>
> p.value<-pchisq(chisq.test.statistic, df, lower.tail = FALSE);
> p.value
[1] 0.7035373
```

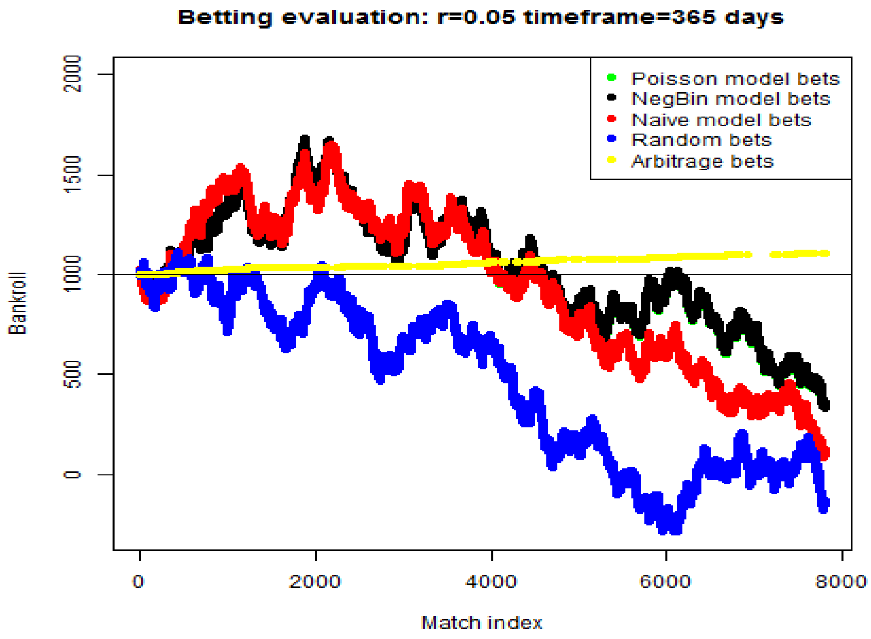
```
#Chi-square independence test

chisq.test(independence.test)

Pearson's Chi-squared test

data:  independence.test
X-squared = 48.551, df = 16, p-value = 3.89e-05
```

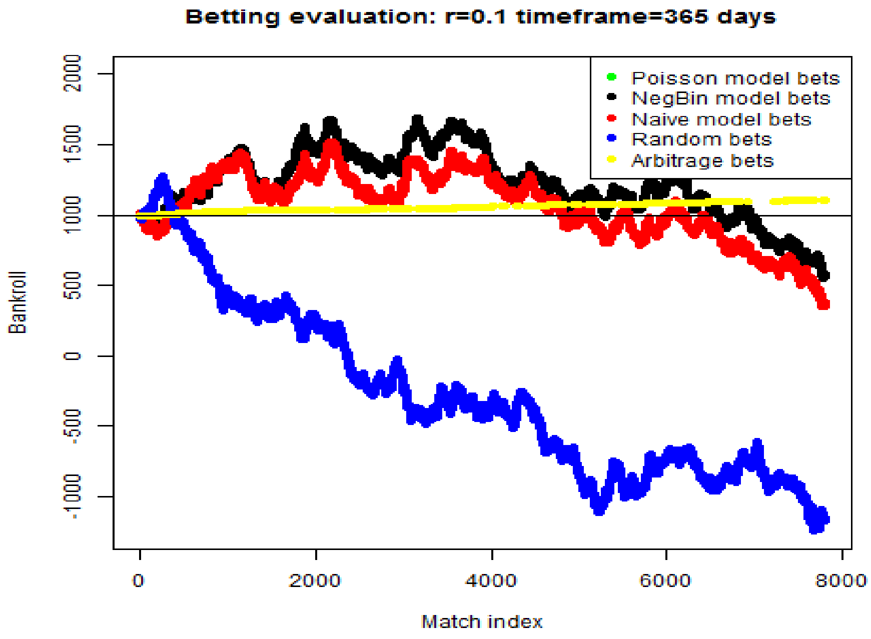
```
> #Goodness-of-fit deviance test for Poisson home goals

> pchisq(model.homegoals$deviance, model.homegoals$df.residual, lower.tail = FALSE)
[1] 0.60505
```

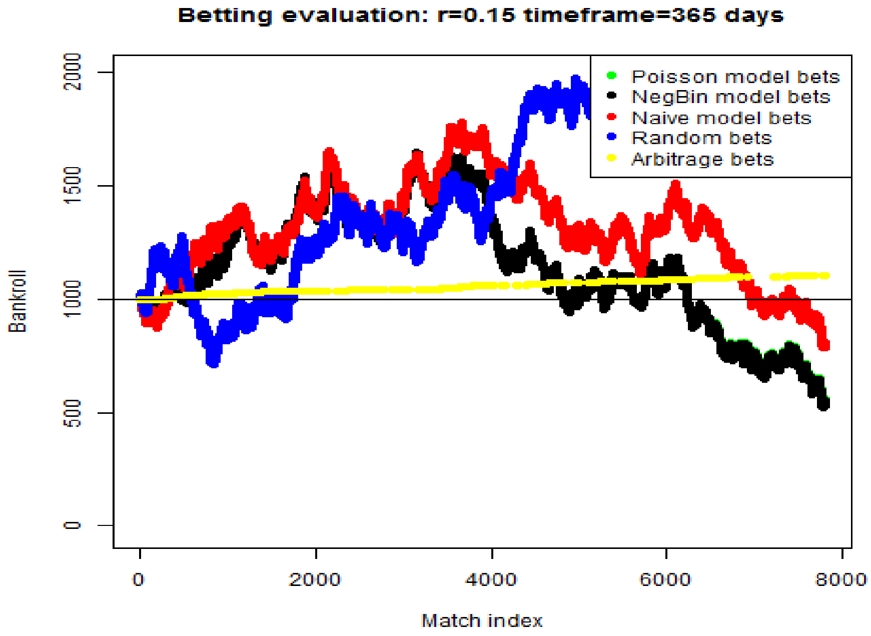
```
> #Goodness-of-fit deviance test for negative binomial home goals

> pchisq(negbin.model.homegoals$deviance,
negbin.model.homegoals$df.residual, lower.tail = FALSE)
[1] 0.6052919
```

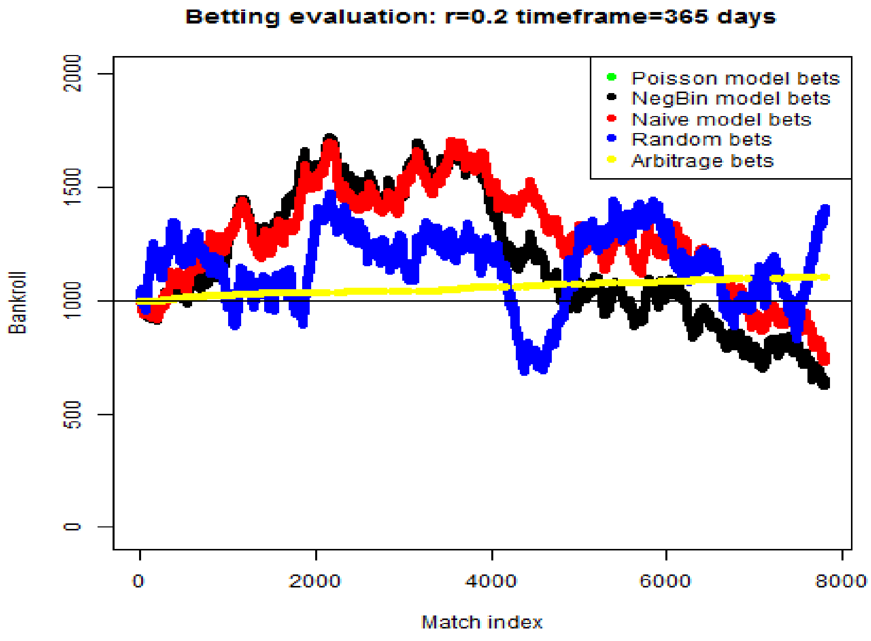
```
> #Goodness-of-fit deviance test for Poisson away goals

> pchisq(poisson.model.awaygoals$deviance,
poisson.model.awaygoals$df.residual, lower.tail = FALSE)
[1] 0.1399714

> #Goodness-of-fit deviance test for negative binomial away goals

> pchisq(negbin.model.awaygoals$deviance,
negbin.model.awaygoals$df.residual, lower.tail = FALSE)
[1] 0.140193
```

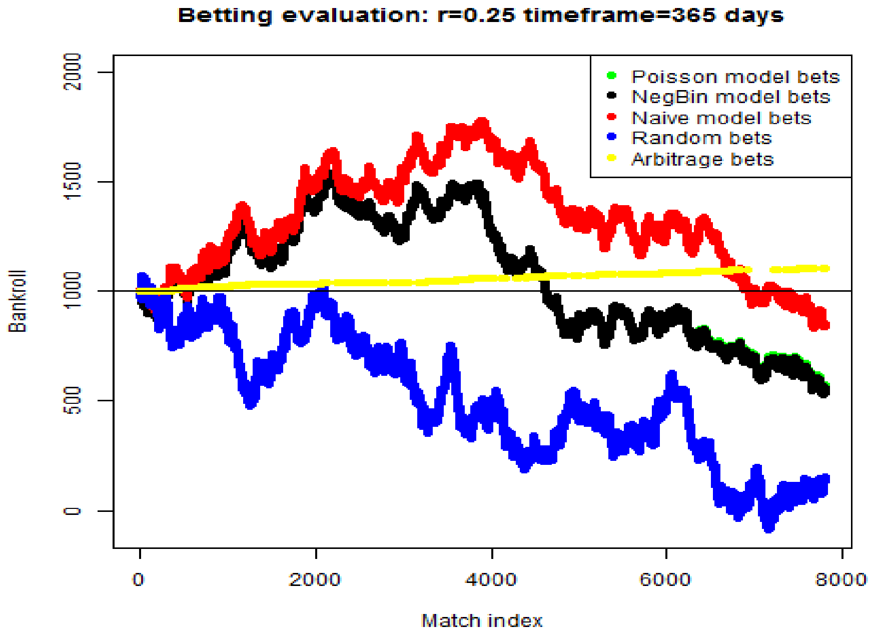
8.2 Evaluation Plots

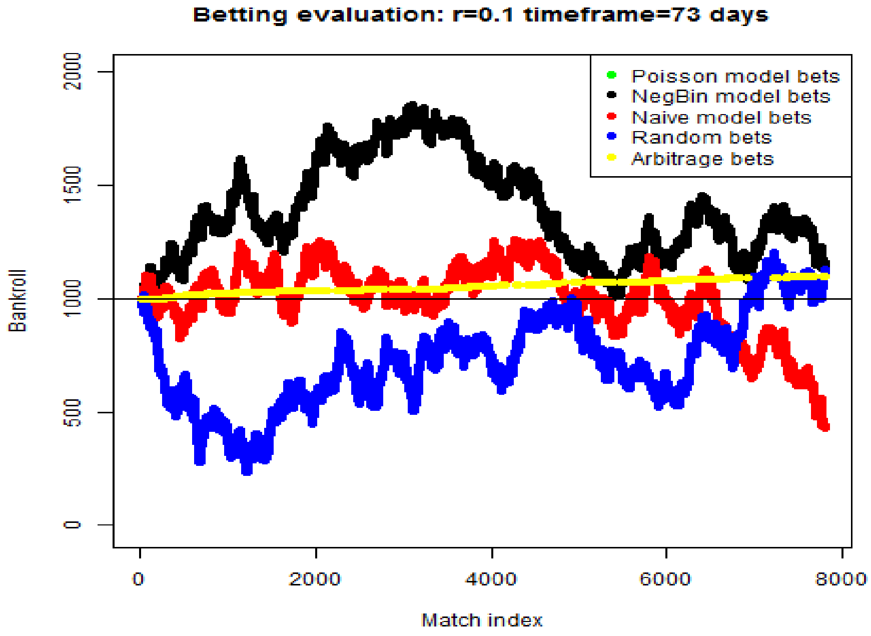


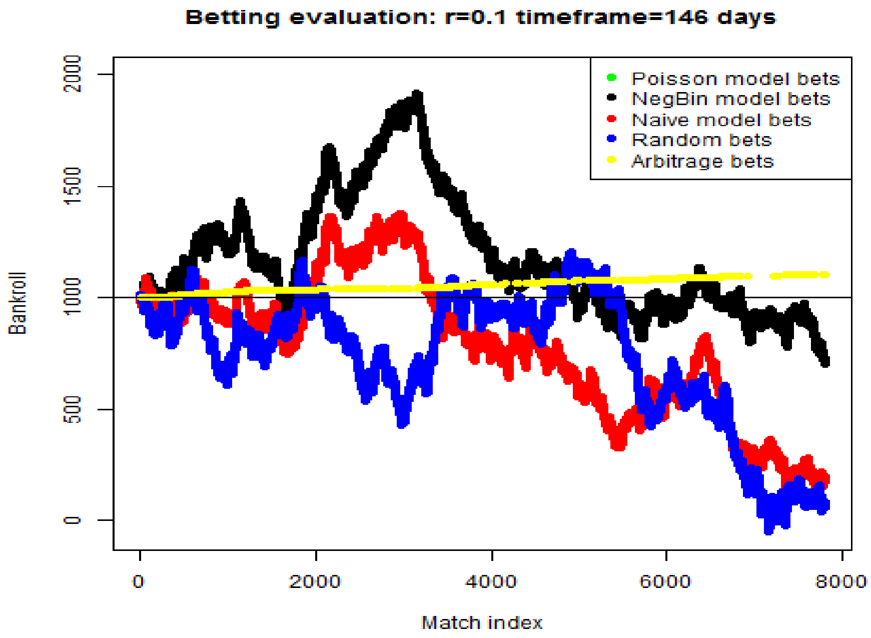


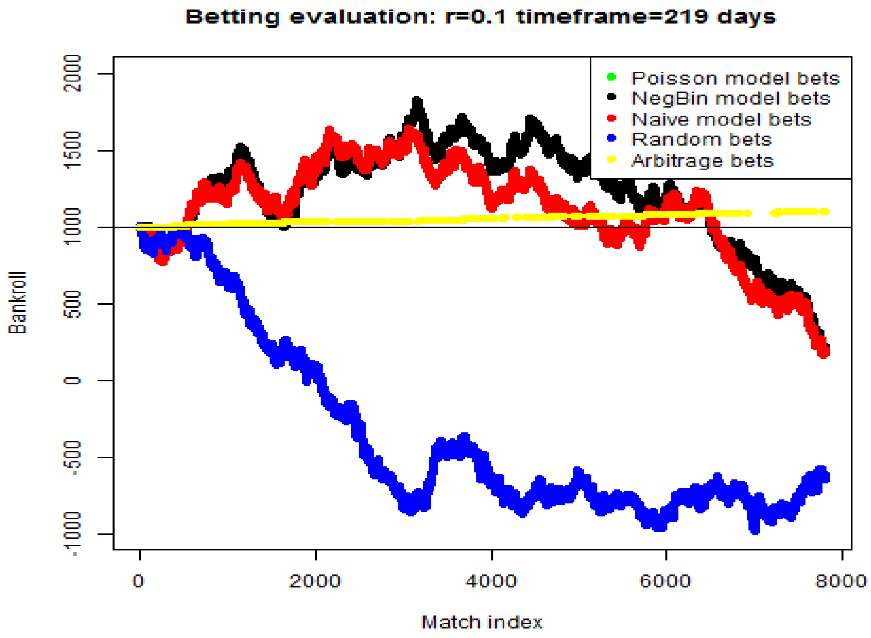

Betting evaluation: r=0.15 timeframe=365 days
● Poisson model bets
● NegBin model bets
● Naive model bets
● Random bets
● Arbitrage bets
Bankroll
0
500
1000
1500
2000
0
2000
4000
6000
8000
Match index

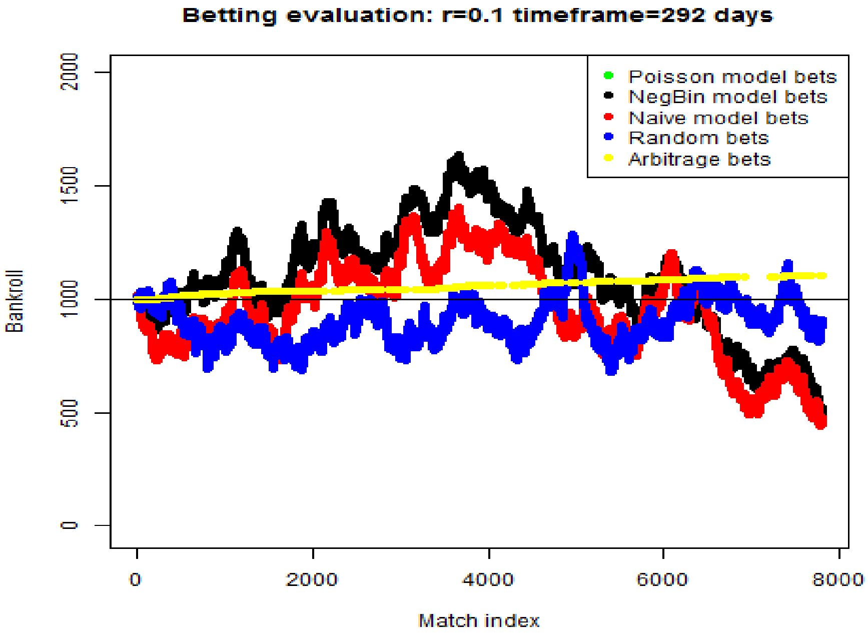

Betting evaluation: r=0.2 timeframe=365 days
● Poisson model bets
● NegBin model bets
● Naive model bets
● Random bets
● Arbitrage bets
Bankroll
0
500
1000
1500
2000
0
2000
4000
6000
8000
Match index

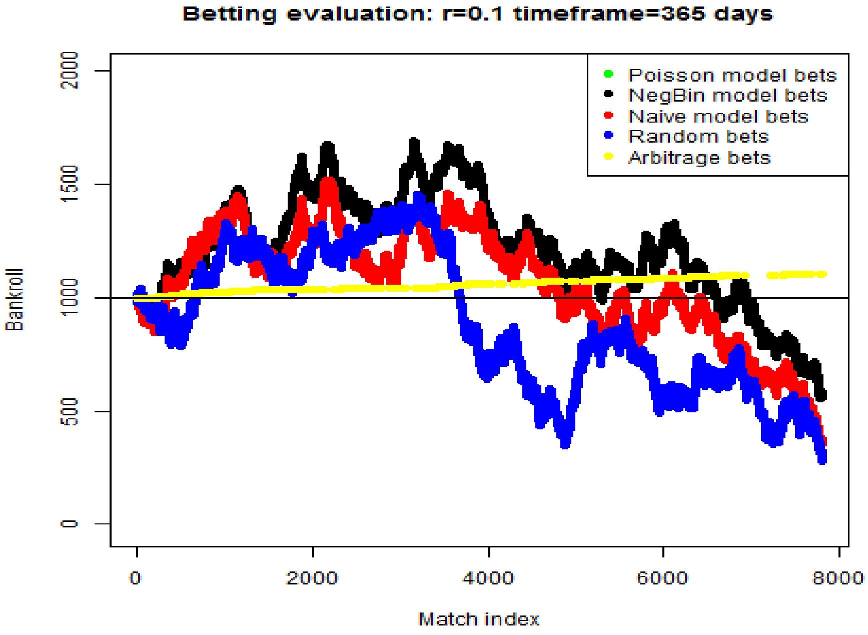

Betting evaluation: r=0.25 timeframe=365 days
Bankroll
0
500
1000
1500
2000
0
2000
4000
6000
8000
Match index
● Poisson model bets
● NegBin model bets
● Naive model bets
● Random bets
● Arbitrage bets

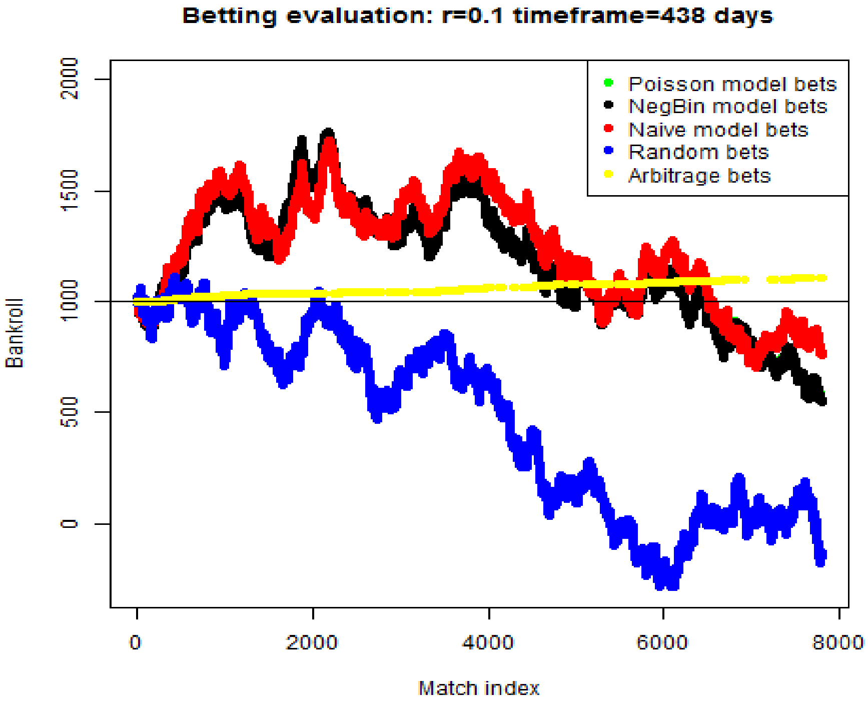

Betting evaluation: r=0.1 timeframe=73 days
Bankroll
0
500
1000
1500
2000
0
2000
4000
6000
8000
Match index
● Poisson model bets
● NegBin model bets
● Naive model bets
● Random bets
● Arbitrage bets

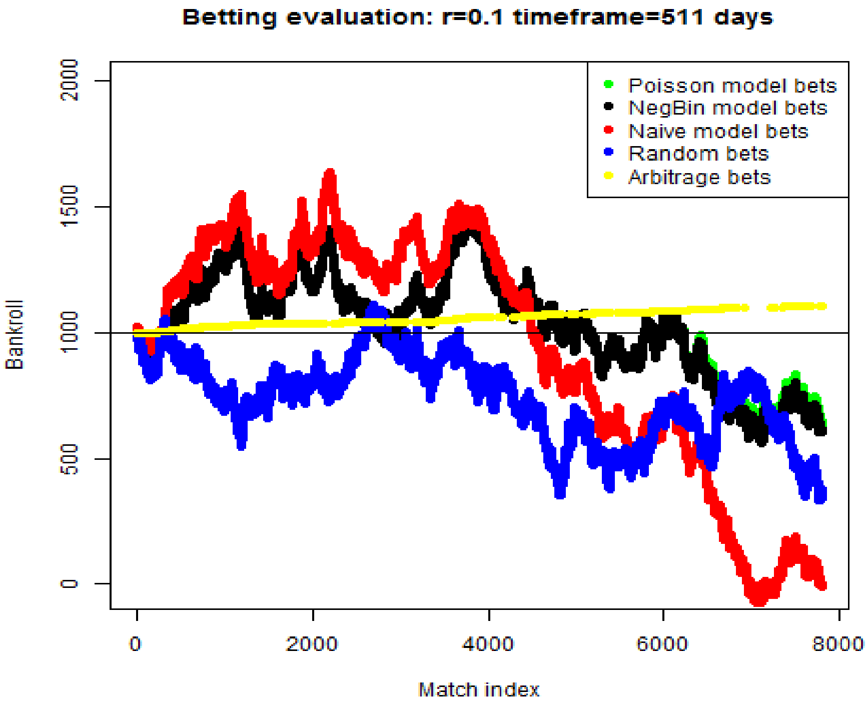

Betting evaluation: r=0.1 timeframe=146 days
Bankroll
0
500
1000
1500
2000
0
2000
4000
6000
8000
Match index
● Poisson model bets
● NegBin model bets
● Naive model bets
● Random bets
● Arbitrage bets

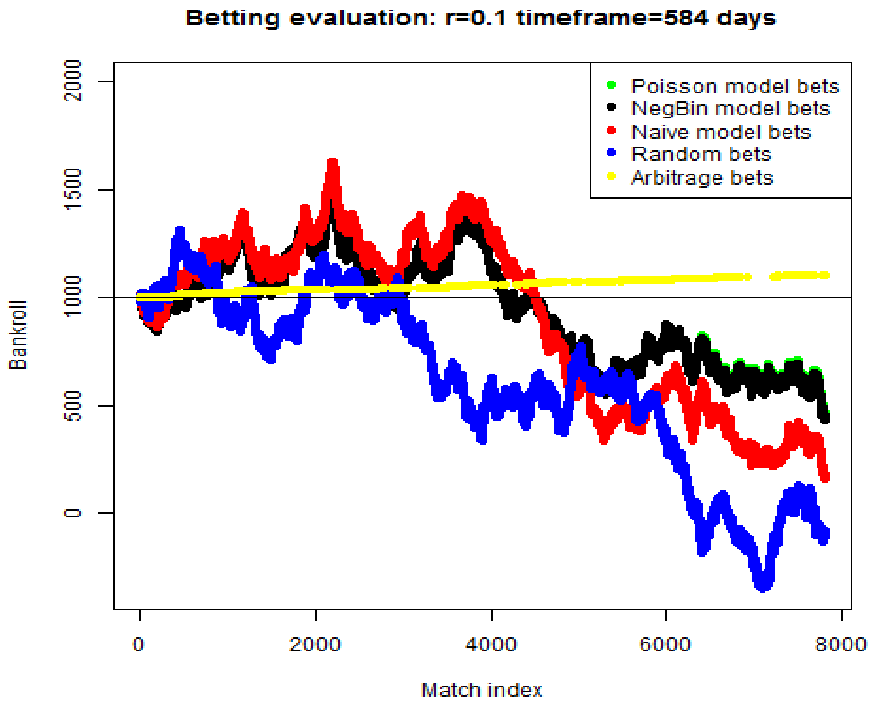

Betting evaluation: r=0.1 timeframe=219 days
Bankroll
-1000
-500
0
500
1000
1500
2000
0
2000
4000
6000
8000
Match index
● Poisson model bets
● NegBin model bets
● Naive model bets
● Random bets
● Arbitrage bets

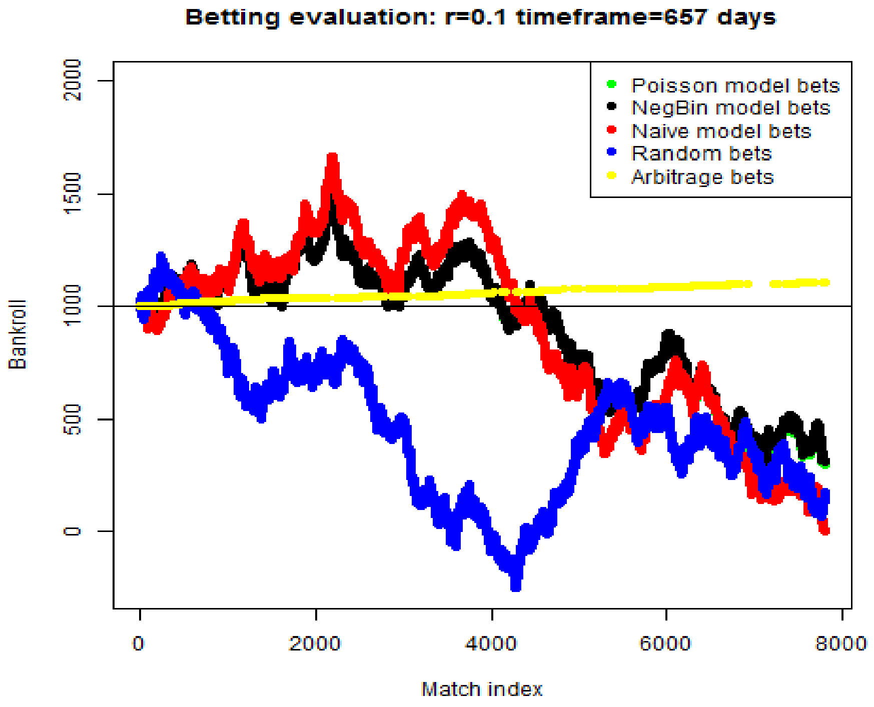

Betting evaluation: r=0.1 timeframe=292 days
Bankroll
0
500
1000
1500
2000
Match index
0
2000
4000
6000
8000
Poisson model bets
NegBin model bets
Naive model bets
Random bets
Arbitrage bets


Betting evaluation: r=0.1 timeframe=365 days
Bankroll
0
500
1000
1500
2000
Match index
0
2000
4000
6000
8000
Poisson model bets
NegBin model bets
Naive model bets
Random bets
Arbitrage bets


Betting evaluation: r=0.1 timeframe=438 days
Poisson model bets
NegBin model bets
Naive model bets
Random bets
Arbitrage bets
Bankroll
0
500
1000
1500
2000
0
2000
4000
6000
8000
Match index


Betting evaluation: r=0.1 timeframe=511 days
Poisson model bets
NegBin model bets
Naive model bets
Random bets
Arbitrage bets
Bankroll
0
500
1000
1500
2000
0
2000
4000
6000
8000
Match index


Betting evaluation: r=0.1 timeframe=584 days
Poisson model bets
NegBin model bets
Naive model bets
Random bets
Arbitrage bets
Bankroll
2000
1500
1000
500
0
0
2000
4000
6000
8000
Match index


Betting evaluation: r=0.1 timeframe=657 days
Poisson model bets
NegBin model bets
Naive model bets
Random bets
Arbitrage bets
Bankroll
2000
1500
1000
500
0
0
2000
4000
6000
8000
Match index