Referee Analysis

Referees have an influential role in any football match - a missed call or bad decision can change the game in an instant. It may be the case that it would be beneficial for managers to have the referee in mind when picking his line-up, for example it might not be best to play a very aggressive defender when the referee is very strict and has a habit of giving out red cards.
We analyse 3 different things:
- Firstly we investigate if some referees tend to be lenient and some strict or if all referees behave in similar ways.
- It seems plausible that that the home crowd might influence the referee. The referee knows a favourable decision for the home team will lead to cheers while being lenient towards the away team causes uproar. We compare statistics for home and away teams to see if referees tend to please the crowd or manage to stay completely unbiased.
- While giving out cards certainly can change change the game, at the end of the day what really matters is the number of goals scored. A lenient referee leads to a free flowing game while a strict referee usually tends to cause lots of starts and stops due to fouls. We investigate if referees have any affect on goals per game.

We analysed data for premier league referees that have refereed at least 50 games over the last 10 years of the league. There are 22 such referees, which for graphing purpose we label as referee 1 to 22.

Referees gave out an average of 3.24 yellow cards a game and 0.16 red cards per game.
M Halsey gave out the fewest yellow cards per game at 2.03 while M Riley gave out the most at 3.83. R Styles was the most likely to send someone off with 0.23 red cards per game while H Webb likes to keep the players on the pitch with only 0.11 reds per game.
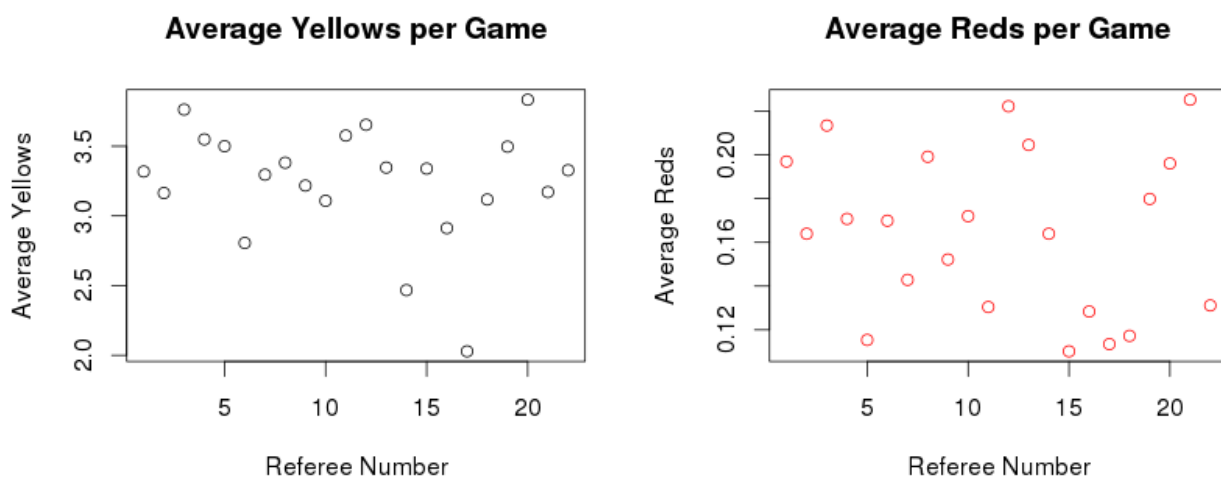
Most referees behave relatively similarly in terms of yellow cards, with the majority of referees giving out somewhere between 3 and 3.5 yellows. However there is more variance in the amount of reds – There are several referees who give approximately 0.12 reds per game and several who give nearly double that amount at 0.20. This is a very substantial difference. This leads us to believe that there certainly is a distinction between different referees and that a manager should keep the referee in mind in terms of their planning.

We now compare the number of cards for home teams and away teams.

**Yellows- Home(blue) vs Away(red)**

Reds- Home(blue) vs Away(red)
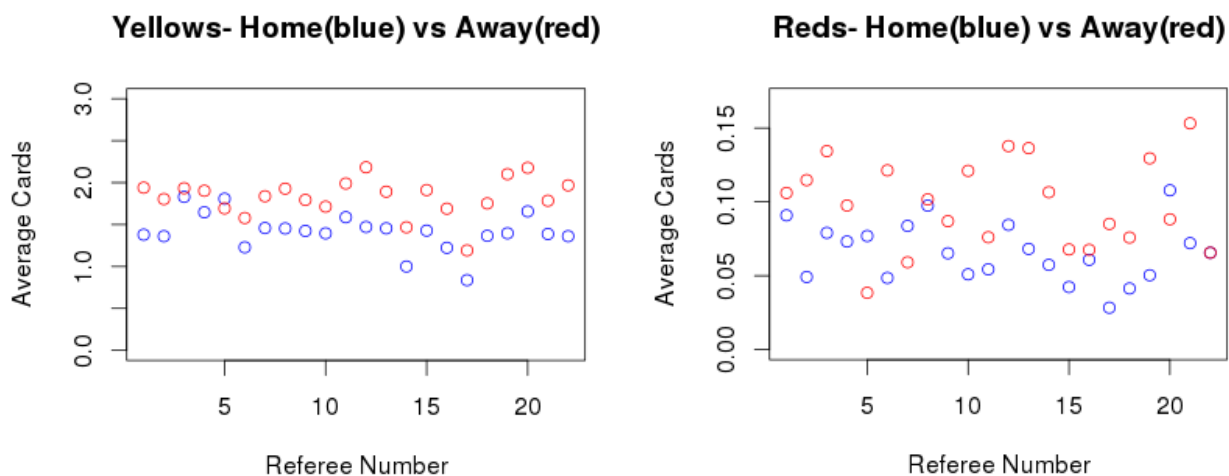
It can be seen that nearly every referee gives more cards to the away team than home team for both yellow and red cards. Referees give an average of 1.42 yellows and 0.066 reds to home players compared to 1.83 yellows and 0.099 reds for away players. This means that referees give an average of 29% more yellow cards to away players and an astonishing 50% more red cards to away players!

Fouls are called on players who are defending so if a team spends large proportion of the game on the defensive end then you would expect them to commit more fouls and hence receive more cards. Since home teams generally tend to spend more time attacking it could be argued that this is the reason for our findings.

However looking at average fouls per game we find that the home team commits 11.46 fouls per game compared to 11.87 for the away team. This difference is negligible. There doesn't seem to be much of an argument for why a foul a home player commits would be any better/worse in severity compared to one an away player commits - hence we deduce that it is primarily the referees decision making that influences the number of cards. This leads us to believe that referees really do get substantially swayed by the crowd! Managers should keep this in mind in their game plan, ensuring that the team is more disciplined in defence for away games.

Finally we look at the influence referees have on the number of goals per game. Our first thought was to simply compare goals per game for each referee:

Average Goals per Game
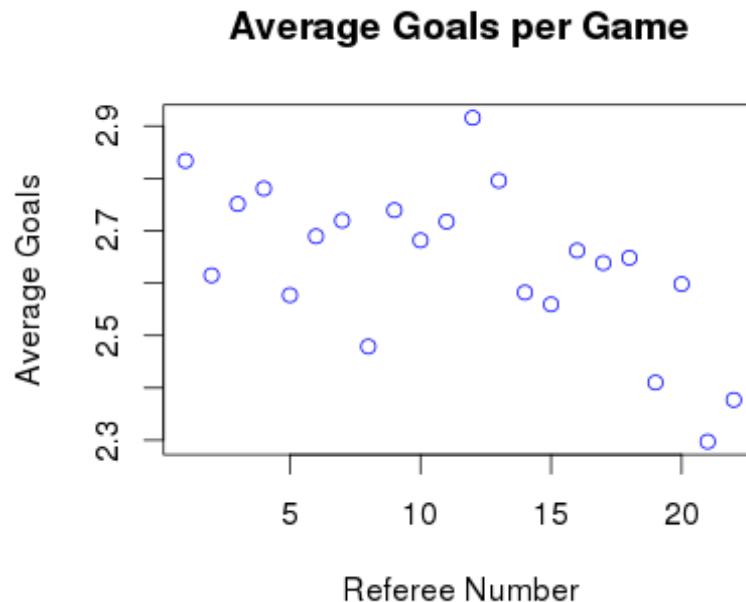
All referees allow somewhere between 2.2 goals and 3 per game. Every goal counts in football so 0.8 of a difference really is substantial. From looking at this alone it would seem to suggest that referees do influence the number of goals scored. However, upon more thought we realise simply comparing average goals per game isn't the best thing to do - it might be that some referees just happen to often referee the teams that are highly offensive and naturally high scoring.
Hence we do the following for each referee:
- Compute the expected goals for a given game based on the teams playing. (This is based on the average goals scored and conceded for the teams involved). Finding the average of this over all games gives us the expected goals per game.
- Compare this with the actual goals per game.

Goals- Actual(blue) vs Expected(red)
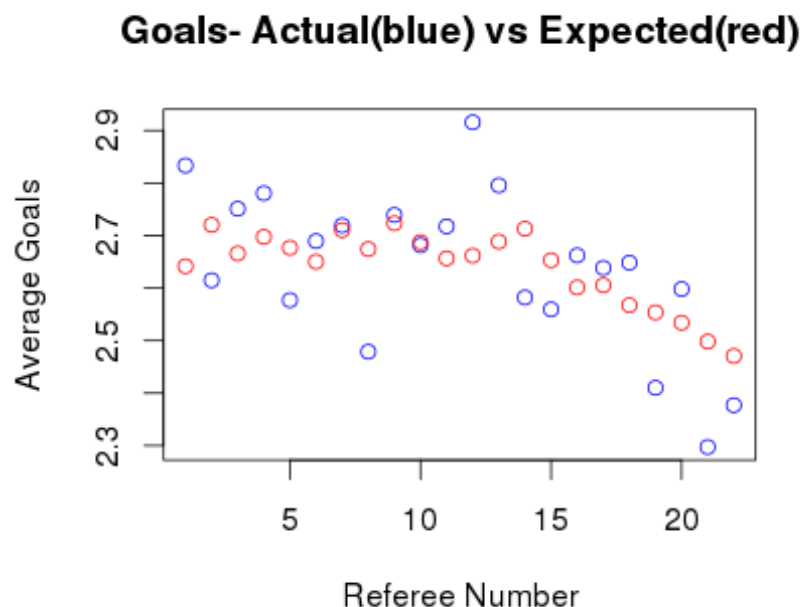
It still is not clear if the referee does influence the game. We would not expect expected goals and average goals to be completely aligned due to the randomness of games. However there are a few referees for which these do differ substantially, for example P Dowd allowed 0.25 more goals than would be expected, while M Atkinson left in 0.20 less.

To investigate we perform a hypothesis test:
-Null Hypothesis: Referees do not have any impact on goals per game
-Alternative Hypothesis: Referees do impact goals per game

First we calculate the total expected goals per game (denote E) and actual goals per game (denote A) for each referee by summing up over all the game they have refereed.
The number of goals in a game is very well approximated by a Poisson model, where the parameter is simply the expected goals per game. We have already calculated expected goals per game for a given match (call this x). Hence we know the goals for this game has an approximate Poisson(x) distribution. We recall the property:

Poisson(x) + Poisson(y) ~ Poisson(x+y)   (when x,y independent)

It is quite a fair assumption to assume goals in different games are independent. If we sum up all the expected goals per game we find the total expected number of goals (Lets denote this by E).
From the additive property it follows that the total goals should have a Poisson(E) distribution.
Since the cumulative distribution function of a Poisson is easy to use and we know the real total number of goals we can calculate the probability that we would get a value at least as extreme as the one we have seen. (Two-sided test)
We calculate this probability for each referee and sort them into order. Assuming the null hypothesis we expect this data to come from a Uniform(0,1) distribution. If this were true, plotting probability vs. the order statistics should approximately give us the straight line x=y.

Probabilities for this Data
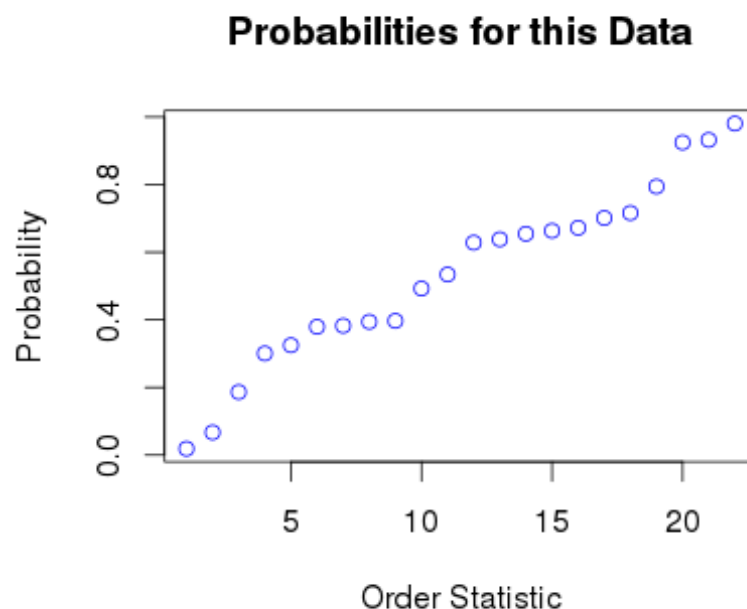
This is close to a straight line (correlation coefficient = 0.9816327). Hence the data fits the null hypothesis very well so we conclude that referees do not have any influence on the number of goals per game.