The Validity of Applying the Colley Method, Massey Method, and Elo Rating on Rankings in National Football League Regular Seasons

Yixuan (King) Zhang
G11 AP statistics section 1
Beijing National Day School, Beijing, China

Proposal

1. Research Question

Based on daily observation and personal understanding of sports, the group thinks that merely ranking teams based on winning percentage cannot tell everything about the power rankings of teams. We learn that there are several ranking methods that might apply. Therefore, we will examine the performance of other methods in giving a final ranking for National Football League (NFL) teams in regular seasons. Potential applicants include the Colley Method, Massey Method, and Elo Rating. Then, the group will analyze the advantages and disadvantages of applying each method, based on their working mechanisms.

Finally, the group would like to modify these rating methods and develop of new one that avoids all potential drawbacks to the best of our ability. This might or might not come true, due to the matter of time.

2. Background Research

In this portion, the group will introduce the mechanisms of each applicant mentioned above.

2.1 Colley Method

To offer readers a clear view of its principle, the group will break down all parts of the Colley Method. Colley first develops a statistic,

$$r_i = \frac{1 + n_{w,i}}{2 + n_{tot,i}}$$

where

$n_{w,i}$ represents the number of winning games for team *i*,

$n_{tot,i}$ represents the number of total games for team *i,*

$n_{l.i}$ represents the number of losing games for team *i.*

This statistic differs from the winning percentage because it adds 1 to the numerator and 2 to the denominator, meaning that the initial value of r is 1/2, and fluctuates around it. Further, $n_{w,i}$ can be transformed into other forms,

$$n_{w,i} = \frac{n_{w,i} - n_{l,i}}{2} + \frac{n_{w,i} + n_{l,i}}{2} = \frac{n_{w,i} - n_{l,i}}{2} + \frac{n_{tot,i}}{2}$$

Since the initial value of r is 1/2 and it fluctuates around 1/2, we can derive that,

$$\frac{n_{tot,i}}{2} \approx \sum_{j \in O_i} r_j$$

Where O_i represents the opponents set for team i. Therefore, we can derive that,

$$n_{w,i} \approx \frac{n_{w,i} - n_{:,i}}{2} + \sum_{j \in O_i} r_j$$

Then,

$$r_i = \frac{1 + \frac{n_{w,i} + n_{l,i}}{2} + \sum_{j \in O_i} r_j}{2 + n_{tot,i}}$$

and we can simply get

$$(2 + n_{tot,i}) r_i - \sum_{j \in O_i} r_j = 1 + \frac{n_{w,i} + n_{l,i}}{2}$$

Furthermore, the system can be written in a linear system, as follows,

$$C\vec{r} = \vec{b}$$

Where $\vec{r}$ is the column vector of all rating of different teams, and $\vec{b}$ is the column vector of

$$1 + \frac{n_{w,i} + n_{l,i}}{2}$$

C, the Colley Matrix, is defined as follows,

$$c_{ii} = 2 + n_{tot,i},$$
$$c_{ij} = -n_{j,i},$$

where $n_{j.i}$ is the times which team j and i played.
Finally, we can solve this linear system to find out the final values of $\vec{r}$.

2.2 Massey Method

The Massey Method is quite similar to the Colley method, but they differ on the grounds that the Massey method take accounts into the point difference between teams. The linear system of the Massey Method can be represented as follows,

$$M\vec{r} = \vec{d}$$

The Massey Method also involves a matrix, M, just like the Colley Matrix. It is defined as follows,

$$m_{ii} = n_{tot,i}$$
$$m_{ji} = m_{ij} = -n_{j,i}$$

The right side of the linear system, $\vec{d}$, is defined as equals the sum of the total point differentials in a season where a win gives a positive differential and a loss, a negative one.

2.3 Elo Rating

Elo Rating is a rating system invented by Hungarian American physics Arpad Elo and is widely applied in chess. However, Elo rating is not widely used in major sports leagues, but since the group found potential availability of this application in sports, we decide to introduce it into our project.

The fundamental concept of Elo Rating is normal distribution. Every player will have a rating of their performance ability, which forms a bell curve. However, not every player has the same ability, so we'd like to find out the difference between two players' ability. And the difference is the portion of overlap of the bell curves of two players. Then, according to Elo, the probability of player A wins when A competes with play B is defined as follows,

$$P(A\ \ wins) = 10^{\frac{R_A - R_B}{400}} \times P(B\ \ wins)$$

And if we further simplify it,

$$P(A\ \ wins) = \frac{1}{1 + 10^{\frac{R_A - R_B}{400}}}$$

After each game, the ratings need to be updated based on how the actual outcomes compare with expected outcomes (probability of winning). The updating formula of ratings can be defined as follows,

$$New\ \ rating = Old\ \ rating + 32 \times (real\ \ score - expected\ \ score)$$

Note that if a player wins a game, they have a real score of one; if they lose, they get 0; if they experience a draw game, they get 1/2. Here the maximum rating that each win brings to a player is 32. It is just what Elo set. There is no special meaning.

After a series of games, all players can get a final rating, and we can rank them based on the rating values.

3. Sampling and Experimental Design

3.1 Variables

The variable in this project is only the results of games (win, lose, or draw, in few cases).

3.2 Type of Study

This is an observational study. All data are collected based on observation (knowing the results of games). There is no control group in this project, so it is not an experiment.

3.3 Data Collection

Results of games in NFL seasons can be found everywhere today due to the development of internet. The main website the group uses is pro-football-reference.com. However, the format arrangement of data is a big deal in this project because there is no pre-made table that shows which team wins the other. Also, the group need to adjust the format of data in order to be fit for computer programming.

3.4 Scope of Inference

The group will randomly select the data of 5 NFL seasons to analyze. These data can be generalized to the population of all NFL seasons.

4. Explanatory Data Analysis

The group will randomly select the data of 5 NFL seasons to analyze. For each season, we will apply the above 3 rating methods and calculate the ratings for each. Then, by certain means, we will compare the rankings generated by these ranking methods with the true ranking (based on winning percentage), and then we will conduct a series of Chi-square goodness-of fit tests to find out how the results of these ranking methods match with the true ranking.

5. Group Task Assignments and Deadlines

5.19 Data collection and arrangement

5.23 Run the codes, calculate the ratings, and conduct the hypothesis tests.

5.25 Finish all the writing

5.26 Final check

6. Data

Week	Winner/tie		Loser/tie	Pts	Pts
1	*Tampa Bay Buccaneers*		Dallas Cowboys	*31*	29
1	*Philadelphia Eagles*	@	Atlanta Falcons	*32*	6
1	*Pittsburgh Steelers*	@	Buffalo Bills	*23*	16
1	*Carolina Panthers*		New York Jets	*19*	14
1	*Cincinnati Bengals*		Minnesota Vikings	*27*	24
1	*Seattle Seahawks*	@	Indianapolis Colts	*28*	16
1	*Arizona Cardinals*	@	Tennessee Titans	*38*	13
1	*San Francisco 49ers*	@	Detroit Lions	*41*	33
1	*Houston Texans*		Jacksonville Jaguars	*37*	21
1	*Los Angeles Chargers*	@	Washington Football Team	*20*	16
1	*Kansas City Chiefs*		Cleveland Browns	*33*	29
1	*Denver Broncos*	@	New York Giants	*27*	13
1	*New Orleans Saints*		Green Bay Packers	*38*	3
1	*Miami Dolphins*	@	New England Patriots	*17*	16
1	*Los Angeles Rams*		Chicago Bears	*34*	14
1	*Las Vegas Raiders*		Baltimore Ravens	*33*	27

The above is the schedule of the first week of NFL 2021 season. This is only about one percent of the whole data set.

7. Reference

[1] W. N. Colley, Colley's Bias Free College Football Ranking Method: The Colley matrix explained, (2002) available at: http://www.colleyrankings.com/method.html.

[2] K. Massey, Statistical models applied to the rating of sports teams, undergraduate honors thesis, Bluefield College, 1997.

[3] Elo, Arpad E. (August 1967). "The Proposed USCF Rating System, Its Development, Theory, and Applications" (PDF). Chess Life. XXII (8): 242–247.

[4] Data source: https://www.pro-football-reference.com/