Predictive Efficacy of a New Association Football League Format in Polish Ekstraklasa

Jan Lasek[1] Marek Gagolewski[2]

[1]Interdisciplinary PhD Studies Program
Institute of Computer Science, Polish Academy of Sciences
janek.lasek@gmail.com

[2]Systems Research Institute, Polish Academy of Sciences
and
Faculty of Mathematics and Information Science, Warsaw University of Technology
gagolews@ibspan.waw.pl

Machine Learning and Data Mining for Sports Analytics
Workshop at ECML & PKDD

Introduction

Domestic football championships can have several different forms. There are changes with respect to the rules according to which

- a champion is selected
- teams are qualified for international cups
- teams are relegated.

In this exposition, we are interested **which league format produces the strongest team as a winner with higher probability**. We will investigate two different league forms.

Different league formats

among countries belonging to UEFA

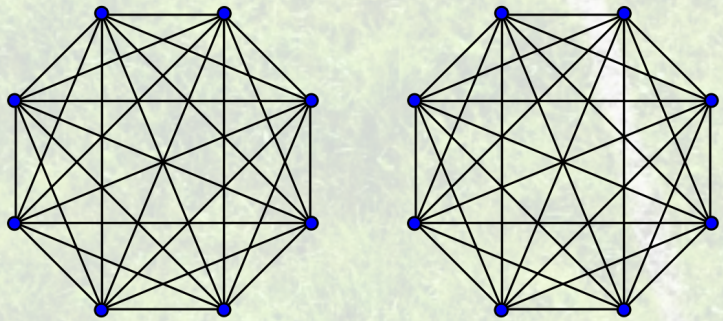
Most prevalent league format is a double round–robin tournament.

- English Premiership, Spanish La Liga, German Bundesliga, Italian Serie A, French Ligue 1, . . . , all operate as a double round–robin tournament
- Polish Ekstraklasa, Belgian Jupiter League, Dutch Eredivisie, Scottish Premiership, Kazakh Premier League operate in different (and diverse) formats

A double round–robin tournament

In a double round–robin tournament with n teams, each team play each other once – home and away. In total $2 \cdot (n-1)$ rounds and $2 \cdot \binom{n}{2}$ games are played.

In case of Polish league with 16 teams at stake, this would give 30 rounds and 240 matches in a season played in this competition format.



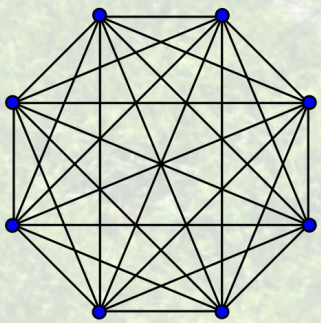
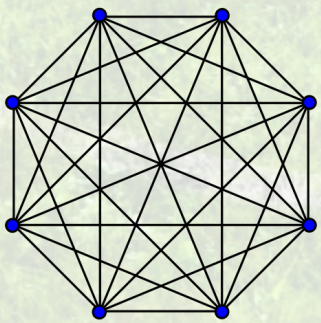
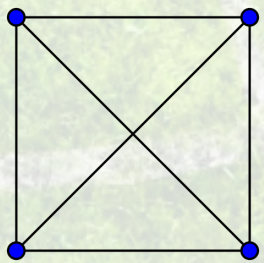
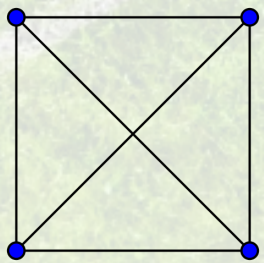
Two–stage league design

The league format currently in force in Polish Ekstraklasa comprises of two stages.

- During the first stage teams compete in *a double round–robin tournament*
- In the second stage, the table is divided into two halves: championship and relegation groups and within each group *a single round–robin tournament* is played

After the first stage **the accumulated points are divided by two** (with possible rounding halves up).

In total $2n + n/2 - 3$ rounds $2 \cdot \binom{n}{2} + 2 \cdot \binom{n/2}{2}$ games are played (with n even). For Ekstraklasa, this yields $30 + 7$ rounds and 296 matches.

Question to you!

Which league format – a regular double-round robin system or the described system currently in force in Polish Ekstraklasa – **produces as the winner the best team with higher probability**? If yes, to what extent one of the formats is superior to the other?

Experiment setup

In the consecutive parts we present the components of our simulation experiment for calculating probability of the strongest team's win. In particular, we discuss:

- game outcome model
- distribution of team ratings
- evaluation metrics

Game outcome model – ordinal logistic regression

Let a team be characterised by a single parameter indicating its strength – a rating r_i. Denoting $d_{ij} = h + r_i - r_j$ (parameter h accounts for the home team advantage), the observed outcome of a match is

$$R_{ij} = \begin{cases} H_{ij} & \text{if } d_{ij} + \epsilon \geq c, \\ D_{ij} & \text{if } d_{ij} + \epsilon \in (-c, c], \\ A_{ij} & \text{if } d_{ij} + \epsilon < -c. \end{cases} \tag{1}$$

Under the assumption that the random component follows the logistic distribution with mean equal to 0 and scale parameter equal to 1, we have

$$\begin{aligned} \mathbb{P}(H_{ij}) &= 1 - \frac{1}{1 + e^{-c+d_{ij}}}, \\ \mathbb{P}(D_{ij}) &= \frac{1}{1 + e^{-c+d_{ij}}} - \frac{1}{1 + e^{c+d_{ij}}}, \\ \mathbb{P}(A_{ij}) &= \frac{1}{1 + e^{c+d_{ij}}}. \end{aligned}$$

Team ratings distribution

We need to decide how to choose rating values for the teams, r_i for $i = 1, \ldots, N$. To this end, we adopt several ratings distributions

- estimated ratings from previous seasons' data
- normal distribution
- Pareto distribution
- exponential distribution

Model calibration

To calibrate the model, we choose dispersion parameter of the given distributions so as to proportion of results (H,D,A) is roughly corresponding to the observed outcomes in European leagues.

For these proportions for 2014/2015 season we have that

- home team wins varies from 40% in Italy up to 53% in Greece
- draws varies from 19% in Scottish Premier League to 31% in Italian *Serie A*
- away team wins varies from Scotland (36%) and the lowest in Greece (22%)

Tournament (evaluation/comparison) metrics

We compare rankings based on **teams' true parameter strength** and **final rankings** produced by a given league design.

The rankings are compared according to the following metrics

- probability that the strongest team wins π_i
- Kendall correlation coefficient τ_i between theoretical ranks and the ranks produced by a league.

In the subsequent slides we presents results of simulation of 100,000 league tournaments.

Results

Kernel density-estimated with different bandwidths.

σ_h	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
π_1	0.44	0.46	0.48	0.51	0.55	0.58	0.61	0.63	0.66	0.67
π_2	**0.47**	**0.49**	**0.51**	**0.54**	**0.56**	**0.61**	**0.64**	**0.66**	**0.68**	**0.70**
τ_1	0.44	0.50	0.56	0.61	0.66	0.69	0.72	0.75	0.77	0.79
τ_2	**0.46**	**0.52**	**0.58**	**0.63**	**0.68**	**0.71**	**0.74**	**0.77**	**0.78**	**0.80**

Normal distributions of ratings with different σ.

σ	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.1**	**1.2**
π_1	0.34	0.42	0.49	0.54	0.58	0.62	0.64	0.67	0.68	0.70
π_2	**0.37**	**0.45**	**0.52**	**0.57**	**0.61**	**0.64**	**0.67**	**0.69**	**0.71**	**0.72**
τ_1	0.47	0.56	0.62	0.67	0.71	0.74	0.76	0.78	0.79	0.81
τ_2	**0.49**	**0.58**	**0.64**	**0.70**	**0.73**	**0.75**	**0.78**	**0.79**	**0.81**	**0.82**

Result cnd.

Exponential distributions with different choices of rate μ.

μ	**0.8**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.2**	**2.4**	**2.6**
π_1	0.81	0.79	0.76	0.74	0.71	0.69	0.66	0.63	0.61	0.59
π_2	**0.83**	**0.81**	**0.79**	**0.76**	**0.73**	**0.70**	**0.68**	**0.65**	**0.63**	**0.60**
τ_1	0.74	0.70	0.66	0.63	0.59	0.56	0.54	0.51	0.49	0.47
τ_2	**0.76**	**0.72**	**0.68**	**0.64**	**0.61**	**0.58**	**0.55**	**0.53**	**0.51**	**0.49**

Pareto distributions with different choices of scale parameter s.

s	**0.25**	**0.3**	**0.35**	**0.4**	**0.45**	**0.5**	**0.55**	**0.6**	**0.65**	**0.7**
π_1	0.63	0.71	0.76	0.80	0.83	0.85	0.87	0.89	0.90	0.91
π_2	**0.66**	**0.72**	**0.78**	**0.81**	**0.84**	**0.86**	**0.88**	**0.90**	**0.91**	**0.92**
τ_1	0.42	0.48	0.53	0.57	0.60	0.63	0.66	0.68	0.70	0.72
τ_2	**0.44**	**0.50**	**0.54**	**0.59**	**0.62**	**0.65**	**0.67**	**0.70**	**0.72**	**0.73**

Conclusions

Based on the result of the experiment we can draw conclusions that the two–stage league format

- has better ability to produce the strongest team as a winner of the league and
- yields a final rankings of the teams that has higher correlation with theoretical teams's strength parameters than the double round–robin tournament.

The new league design beats the round–robin tournament according to these criteria though not by a large margin: the differences are of order 0.01, but still significant. The more the samples, the better the estimates.

Limitations for the study

However, there are a couple of limitations of our study.

- First of all, changes in the rules of play influence the game
- Team's strength parameters are kept constant – the model is insensitive to team's shape fluctuations, injuries, scheduled cup games, etc.

Further work

on bringing model closer to reality

Primarily, we want to extend the basic model as a part of further work.

- One wants to relax the assumption of static modelling of team ratings with the use of a dynamic model, e.g., an autoregressive process $r_i^{(t+1)} = r_i^{(t)} + u^{(t+1)}$
- Perhaps other factors influencing current team strength (rating) can be incorporated
- Presumably, teams have different attitude (motivation) toward the two individual league stages – this factor could be incorporated into the extended model.

That is all

Thank you!

Bibliography

- Appleton, D.R.: May the Best Man Win? Journal of the Royal Statistical Society: Series C (The Statistician), Vol. 44, No. 4, pp. 529–538 (1995)
- Ryvkin, D.: The Selection Efficiency of Tournaments. European Journal of Operational Research, Vol. 206, No. 3, pp. 667–675 (2010)
- Scarf, P., Yusof, M.M., Bilbao, M.: A Numerical Study of Designs for Sporting Contests. European Journal of Operational Research, Vol. 198, No. 1, pp. 190–198 (2009)
- Szymanski, S.: The Economic Design of Sporting Contests. Journal of Economic Literature. Vol. 41, No. 4, pp. 1137–1187 (2003)
- Borcas I. and Carrillo, J.D: Do the "The 3–Point–Victory" and "Golden Goal" Rules Make Soccer More Exciting? Journal of Sports Economics, Vol. 5, No. 2, pp. 169–185 (2004)