Applying survival analysis to model time to score in football – the case of the UEFA European Championship 2016

Dušan Mundar
Faculty of organization and informatics, University of Zagreb
Pavlinska 2, Varaždin, Croatia
E-mail: dusan.mundjar@foi.hr

Diana Šimić
Faculty of organization and informatics, University of Zagreb
Pavlinska 2, Varaždin, Croatia
E-mail: diana.simic@foi.hr

Abstract
Operational research (OR), considered as a branch of mathematics with primary focus on solving business problems, can be applied in many other areas. OR methods are used in sports management for more than fifty years. On the other hand, it is difficult to determine where OR ends and another discipline, for example statistics, begins. Generally, it’s considered that operational research has a relation to decision-making. Analysis in our paper is intended to be used as a decision making tool.

We are interested in the dynamics of the goal scoring rate during the match. Time between scores is modeled using Cox semiparametric regression on game characteristics (where match breaks, halftime and end of game cause censoring). Data on European Football Championships EURO 2012 and EURO 2016 are used in the paper. If number of goals in a match was a realization of a Poisson process with a fixed goal scoring rate, then time to score would be exponentially distributed. We compare Cox semiparametric model where the goal rate changes during the game with the model of exponential distribution with a fixed goal rate. Cox regression enables assessment of association between game characteristics (e.g. score difference or time to end) and risk of scoring. Results of the analysis can be used as a decision making tool for in-game changes in tactics (for example player substitution).

Keywords: Cox model, Football, Goals, In-game tactics, Survival analysis, Sports analytics