Introduction

FRITS SPIEKSMA

What makes a good scientific paper in sports? Of course, some ingredients are indispensable: sound methodological approach, new insights, lucid writing... But to top it off, one needs a good story; a story that captures the imagination, a story we can relate to, a story with which we are (perhaps vaguely) familiar. Every sport has good stories—a sport cannot survive without them.

The six papers presented in this section represent research in specific sports that are not covered by the other sections of the anthology; these papers address five different sports. In addition to representing research on a wide variety of sports, the types of questions considered in these papers vary quite a bit. However, the papers in this section also have something in common: they each tell a good story. From the famous Holyfield versus Lewis boxing match to the continuing debate on the hot-hand effect, each of the papers in this section offers an interesting story that expands our understanding. Here, we introduce each of these papers and identify their innovative contributions.

Zaman (2001) addresses a well-known hockey issue known as the Pull-the-Goalie (PtG) problem: at what moment should the coach decide to pull the goalie (i.e., to change the goalkeeper for a regular skater) in order to maximize the chance to score a tying goal when trailing by one (or two) goals late in the game? Research on this question was initiated by Morrison (1976), and a recent paper on this topic has been published by Beaudoin and Swartz (2010). Most approaches to the PtG problem are based on the arrival of goals being modelled as a Poisson process, thereby assuming a constant probability over time of a goal occurring. Then, pulling the goalie increases the scoring rate of the team that pulls its goalie (as well the scoring rate of the opposing team!). In contrast to these traditional approaches, Zaman proposes a Markov chain model that includes the location of the puck in the field as a state. There are seven states: Goal A, Shot A, Zone B, Neutral, Zone A, Shot B, and Goal B. One advantage of this point of view is that the state of the game has an impact on the decision when to pull the goalie. Indeed, one should be more careful when the puck is in one's defensive zone compared to when the puck is in the offensive zone. It turns out that, pulling the goalie with 5–8 minutes left to play is the right decision when trailing by one goal depending upon the location of the puck. This outcome is consistent with outcomes from traditional approaches, and also shows that in practice coaches are too conservative.

The second paper (Dawson and Magee, 2001) also deals with the National Hockey League (NHL). Can we reliably predict performance of NHL players as a function of the player's position (goaltenders, defensemen, forwards), year, and overall rank in the draft (or pick number)? Performance is measured by the number of NHL games played. A recent contribution attributing a value to each pick number is found in Hurley *et al.* (2012). Dawson and Magee focus on three questions: (i) What is the expected difference in number of games played between a player drafted as the i^{th} pick, versus a player drafted as the j^{th} pick (for any i,j)? (ii) How well does a franchise draft its players? (iii) To what extent are European players undervalued in the draft? The authors propose a multiplicative model, and describe a so-called weighted pool-adjacent-violators algorithm to arrive at the estimates. They find for each of the three positions a relation between pick number and number of games played. They also find substantial variation among franchises with respect to the difference between the actual and predicted

number of games played by the franchises' draftees. Finally, it is found that in a particular period (1988–1992), European draftees played more matches than expected.

One of the most intensely debated issues in sport statistics is the "hot-hand" question: are success probabilities of consecutive attempts independent? Clearly, this question turns up in many different sports: basketball, baseball, golf, darts, tennis, volleyball, to name a few; see Bar-Eli *et al.* (2006) for an overview of research on this topic. Some papers find the effect, others do not. In particular, an influential study in basketball by Gilovich *et al.* (1985) finds no evidence of the effect. The third paper (Dorsey-Palamateer and Smith, 2004) addresses the presence of the hot hand in bowling. It is argued that bowling has a number of properties that make it attractive for testing the existence of a hot hand effect. Indeed, the absence of interplay of opponents, the almost identical circumstances prior to each attempt, and the regular time-intervals between throws, make bowling a sport well-suited for studying the presence of the hot hand effect. By categorizing each throw as a strike or a non-strike, a binomial model can be used to model that the success probability is stationary, and that throws are independent: player i has a probability p_i to throw a strike. Analyzing the data reveals that a bowler's success probabilities are not independent of previous trials. More specifically, success breeds success, as the authors say; most bowlers have a higher strike proportion after j consecutive strikes than after j consecutive non-strikes, and this effect increases with the value for j.

Who won the boxing fight? This innocent-looking question is not always so easy to answer, as those who watched the match between Evander Holyfield and Lennox Lewis on March 13, 1999 may remember. This particular match was called a draw, even though many observers felt Lewis had the upper hand. Lee *et al.* (2002) looks at this fight and the September 1999 fight between Oscar de la Hoya and Felix Trinidad with respect to interrater agreement. A boxing match is divided into 12 rounds with each round scored individually by three judges who cannot communicate with each other. An individual scorecard then consists of a sequence of 12 symbols, one for each round, indicating whether a round was won by one opponent, by the other opponent, or was called even. The paper shows that it is possible to compute the probability that a particular scorecard is observed, and thereby offer a way to identify and quantify the variance in scoring behavior of different judges, both official ones and unofficial ones. Of course, this gives the opportunity to assess judges' behavior (see also Balmer *et al.* (2005) who focus on home advantage). The authors further argue that unfamiliarity with the precise scoring rules and the perhaps not well-understood inherent stochastic nature of scoring a boxing fight may cause the perceptual difference between general opinion and the actual outcome of a fight. It also may result in bad advice to a boxer while the match is in progress, as exemplified by the match between Oscar de la Hoya and Felix Trinidad.

More and more, the National Basketball Association (NBA) drafts players who are not seniors in college, but instead sophomores, juniors, or freshmen. Thus, a player in college, aspiring to play in the NBA is confronted with the difficult of decision whether to stay in college for another year or enter the draft. There are obvious pros and cons for each alternative. The fifth paper in this section (Bishop and Gajewski, 2004) provides several examples indicating there is no single good choice. The authors define a statistic (the "bling" number), which assesses a player's potential and can be seen as representing the draft likelihood. This number is computed using first principal components analysis and next performing a logistic regression. The principal components analysis identifies from all criteria two major factors, namely game (minutes played, points made, rebounds, and so on) and body (height, weight, blocks). The logistic regression was then used to predict whether a player was drafted. To balance the error

rate among players predicted to be drafted and players predicted not to be drafted, a cutoff probability of 0.20 was chosen, leading to 64 players that were predicted to be drafted in the 2001 season. Individual analysis of some of these athletes reveals that other factors (heart for the game, leadership) also influence drafting decisions.

In the sixth paper in this section, Marcus (2001) proposes a new table tennis rating system. Notice that a rate represents the strength of a player whereas a ranking system only produces an ordinal list of players. Probably, the best-known rating system in sports is the chess rating system (see e.g., Glickman (1999)). In table tennis, due to the large variety in player characteristics (in number of matches played per year), a Bayesian model is deemed appropriate; such a model treats a player's strength as uncertain and uses conditioning to incorporate new results. After a precise and clear description of the current system in use, the new system is explained and the two systems are compared. The paper argues convincingly that a rating system should not use scores of matches to rate players. In the new approach, the strength of each player is modelled as a random variable. The paper describes how to initialize this strength, how strength evolves over time (by adding a random walk), and how new match results are processed to arrive at adjusted strengths. Resolving this latter issue can be computationally intensive, and the paper discusses ways to speed up the algorithm.

REFERENCES

Balmer, N., Nevill, A., and Lane, A. (2005). Do Judges Enhance Home Advantage in European Championship Boxing? *Journal of Sports Sciences*, 23, 409–416.

Bar-Eli, M., Avugos, S., and Raab, M. (2006). Twenty Years of "Hot Hand" Research: Review and Critique. *Psychology of Sport and Exercise*, 7, 525–553.

Beaudoin, D. and Swartz, T.B. (2010). Strategies for Pulling the Goalie in Hockey. *The American Statistician*, 64, 197–204.

Bishop, T. and Gajewski, B. (2004). Drafting a Career in Sports: Determining Underclassmen College Player's Stock in the NBA Draft. *Chance*, 17, 9–12.

Dawson, D. and Magee, L. (2001). The National Hockey League Draft, 1969–1995. *The American Statistician*, 55, 194–199.

Dorsey-Palmateer, R. and Smith, G. (2004). Bowlers' Hot Hands. *The American Statistician*, 58, 38–45.

Gilovich, T., Vallone, R., and Tversky A. (1985). The Hot Hand in Basketball: on the Misperception of Random Sequences. *Cognitive Psychology*, 17, 295–314.

Glickman, M. (1999). Chess Rating Systems. *American Chess Journal*, 3, 59–102.

Hurley, W., Pavlov, A., and Andrews, W. (2012). A Valuation Model for NFL and NHL Draft Positions. *Journal of the Operational Research Society* 63, 890–898.

Lee, H., Cork, D., and Algranati, D. (2002). Did Lennox beat Evander Holyfield? Methods for Analyzing Small Sample Interrater Agreement Problems. *The Statistician*, 51, 129–146.

Marcus, D. (2001). New Table-Tennis Rating System. *Journal of the Royal Statistical Society*, Series D, 50, 191–208.

Morrison, D.G. (1976). On the Optimal Time to Pull the Goalie: A Poisson Model Applied to a Common Strategy Used in Ice Hockey. *TIMS Studies in Management Sciences, Vol. 4*, North Holland, New York, 137–144.

Zaman, Z. (2001). Coach Markov Pulls Goalie Poisson. *Chance*, 14, 31–35.