Physica A 464 (2016) 182–190

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

The Poisson model limits in NBA basketball: Complexity in team sports

Juan Manuel Martín-González [a], Yves de Saá Guerra [a,d,*], Juan Manuel García-Manso [a], Enrique Arriaza [b], Teresa Valverde-Estévez [c]

[a] Departamento de Física, Universidad de Las Palmas de Gran Canaria, Edificio de Ciencias Básicas, Campus Tafira, 35017, Las Palmas de Gran Canaria, Spain

[b] Centro de Estudios Avanzados, Universidad de Playa Ancha, Valparaíso, Chile

[c] Facultad de Ciencias de la Actividad Física, Universidad Católica de Valencia, Spain

[d] Drago. Centro de educación física, investigación y alto rendimiento, Las Palmas de Gran Canaria, Spain

HIGHLIGHTS

- We proved that basketball scoring does not follow a unique distribution.
- Poisson model fails in some game times, where Power Laws show up.
- The result of the competed basketball games depends on the last minute of the game.
- Events likelihood with dt = 1 s is the highest in last minute; free throws mainly.
- Teams play in order to reach last minute. Basketball behaves as Red Queen Hypothesis.

ARTICLE INFO

Article history:
Received 21 January 2016
Received in revised form 9 June 2016
Available online 25 July 2016

Keywords:
Basketball
Non-linear complex systems
Poisson distribution
Power Law
NBA

ABSTRACT

Team sports are frequently studied by researchers. There is presumption that scoring in basketball is a random process and that can be described using the Poisson Model. Basketball is a collaboration-opposition sport, where the non-linear local interactions among players are reflected in the evolution of the score that ultimately determines the winner. In the NBA, the outcomes of close games are often decided in the last minute, where fouls play a main role. We examined 6130 NBA games in order to analyze the time intervals between baskets and scoring dynamics. Most numbers of baskets (n) over a time interval (ΔT) follow a Poisson distribution, but some (e.g., $\Delta T = 10$ s, $n > 3$) behave as a Power Law. The Poisson distribution includes most baskets in any game, in most game situations, but in close games in the last minute, the numbers of events are distributed following a Power Law. The number of events can be adjusted by a mixture of two distributions. In close games, both teams try to maintain their advantage solely in order to reach the last minute: a completely different game. For this reason, we propose to use the Poisson model as a reference. The complex dynamics will emerge from the limits of this model.

* Corresponding author.
E-mail address: yvesdesaa@gmail.com (Y. de Saá Guerra).

http://dx.doi.org/10.1016/j.physa.2016.07.028

1. Introduction

There have been many attempts to understand sports phenomena through the modern theory of Non-linear Complex Systems; systems which involve large numbers of interacting agents [1–6]; sport, in turn, provides a rich laboratory in which to study competitive behavior in a well-defined way [7].

In the study of team sports, most researchers have focused on classical statistics, based on the Poisson model, negative binomial, extreme events, or random walk processes [8,9,5,7,10,11] with good results in basketball and soccer. Statistical models more related to complexity such as Power Laws, q-statistics, etc. have also been successfully used [12–14], as well as competition dynamics [2,12] and game theory [13], particularly in basketball. This helps us to understand complexity, and shed light on the above mentioned concepts, as well as improve our understanding of these sports.

Unlike other sports such as soccer, where scoring can be considered a rare event and a 0–0 result is a common result; the score in basketball is a parameter which provides continuous, almost instantaneous, information that allows an evaluation of the degree of self-organization of a team, or actions taken in each case. In addition, there are differences in the number of points obtained from a basket: 1 (fouls, free throws), 2, or 3 points. The kind of point scored has not the same importance in different moments of the game [10].

A priori, scoring in basketball can be considered a random process, meaning that it is memoryless [7,11,12]. Therefore, the process must follow a Poisson distribution, in which events are random and independent. Violation of this independence leads to Pareto or Power Law distributions [14,15].

A basketball game is a competitive event. During the game, the degree of competitiveness can fluctuate; hence the intensity of competition may be much higher at certain times of the game time than others. This is manifested in the fight for the lead during the game time [11]. In the final moments of most close games, one team has a small advantage over the other and tries to maintain it over the time remaining. The team with the lower score will try to wipe out this difference. The point is that it is in the final moments of the game that uncertainty (the officials, fatigue, injury, wrong decisions, etc.) can be the key. In these moments when effective management and decision making are most necessary, the information among all elements of the system and the interrelations between them should flow more easily. The complexity of the game hence, may increase, which in some way is reflected in the numbers and types of points that are achieved. The struggle to regain the relative advantage in a complex environment, which seems to have some resemblance to a Red Queen's race, has many solutions: try to keep possession, to force the opposition to shoot, take quick fouls, etc. Thus we can expect an avalanche phenomenon in the number of points achieved, affecting the tail of the distribution and a substantial increase in probability of Power Law distributed phenomena.

In this study, our goal is to use the framework of Poisson random processes and their limits in order to find out the extent to which the Poisson process fails in certain situations and in some extreme situations. Moreover, we try to find complex patterns through scaling analysis or Power Laws; or even regularities to improve the understanding of this sport within the completely random framework of the Poisson model. Hence, in this paper the idea is to use two basic frameworks of reference in order to analyze point scoring in basketball: Poisson (or Negative Binomial) and Power Law distributions. Our interest is not to find and discuss statistical models that best fit each situation, but rather to approach the problem in a more qualitative way. Therefore, in order not to obscure the analysis and discussion of the results, which are addressed towards both complexity and sport professionals, we consider it preferable not to use more sophisticated statistical models.

2. Methods

In this paper we focus on the consecutive points scored by any of either team in each game. As a sample for this analysis we used five NBA regular seasons (i.e., no play-offs), a total of 6130 games that we considered adequate for analysis over short time intervals. We excluded overtimes in order to analyze a homogeneous sample which contains all kinds of games with several different competitive scenarios. Overtimes and play-offs correspond to games with subtly different dynamics that would bias our analysis.

In the following we refer any kind of point scored (1, 2, or 3) by either team as an **event** or **basket**. ΔT represents the predetermined time interval with which we analyze the game; meaning the number of events that took place in that time interval. On the other hand, dt represents the elapsed time between events and indeed is variable.

2.1. Poisson distribution

A game can be considered as an arrival process; over a time interval (ΔT) a number n of field goals (baskets) are scored. The Poisson distribution presents the well-known expression:

$$p(x; \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$$

where $p(x, \lambda)$ is the probability of observing x events in a specific time period ΔT, depending on a parameter λ, which has a precise physical meaning: the mean number of events per time interval, $\lambda = \mu * \Delta T$ with μ being the number of events per second. A Poisson process is characterized by two features: