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 $\mathrm{d}t = 1$ 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, $\mathrm{d}t$ 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:

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

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:

1. Index of Dispersion (ratio of the variance and the mean) $ID = 1$ and the probability of no occurrence is $p(x = 0, \lambda) = e^{-\lambda}$, which determines the number of zeros.

   When the Index of Dispersion (ID) is lower than 1, the process is considered to be under-dispersed, which means that its probability distribution is more clustered around the mean making it more predictable. When the variance is greater than the mean, $ID > 1$, the distribution is considered over-dispersed, indicating larger data dispersion. This case points out the possible existence of clusters in data. We highlight two causes that may be responsible for $ID > 1$; the problem of a long tail where the distribution tail is higher than that predicted by the Poisson model, and can be described by a negative binomial distribution that, for non-integer values of the parameter r, is calculated using the gamma function Γ.

$$p(x; r, p) = \frac{\Gamma(r + x)}{\Gamma(r)\Gamma(x + 1)} p^r (1 - p)^x.$$

   The other cause is when there is an excess of zeros, $p(x = 0, \lambda) > e^{-\lambda}$ or both.
2. The inter-event time (dt) or time between arrivals, follows an exponential distribution $p(dt) = \mu e^{-\mu dt}$, which depends on a single parameter μ and is easily detectable in a semi-log plot. This indicates that a Poisson process has no memory; the probability of an event is unchanged by the occurrence of a prior event.

2.2. Scaling analysis

The Scaling Analysis, *Power Law* distributions (PL) play important roles in describing nonlinear complex systems, and are often used to describe many natural and social phenomena. Scaling or *Power Law* relationships arise commonly as probability or frequency-size distributions, and are characterized by the form $f(x) = Cx^{-\alpha}$ where C is a constant [16]. This function is linear when it is plotted on a log–log scale, and the slope of the resultant straight line gives an estimation of the scale exponent α. Scale-invariance is another characteristic associated with *Power Law* distributions, [17,18], thus such phenomena have the same statistical properties at any scale and are not associated with a particular one. The presence of *Power Laws* has also been suggested to be the signature of systems that show Self-Organized Criticality [19], characterized by avalanche phenomena.

Power Laws are very important because they can reveal underlying regularities in the dynamics of a system. However, the presence of a Power Law by itself is insufficient to specify the mechanism that generates it. One important limitation of this tool is the behavior at the tail of the Power Law distribution where the greatest fluctuations are found [20]. Furthermore, in recent years, improved statistical tests have provided strong evidence for scaling laws over substantial (although limited) ranges of scales [21,22]. Despite these difficulties, there are some advantages in the application of Power Laws for the purposes of diagnoses, characterizations, and even predictions. This allows the comparison of analogous phenomena and the characterization of regions over similar environments. For these purposes we do not need to have absolute certainty that an empirical data set follows a Power Law.

2.3. Close games and non-close games

In a recent paper [10], we noted that final point difference distributions seem to follow several Power Laws. The appearance of these Power Laws helps us to classify the entire sample in **close games** (**CG**) and **non-close games** (**NCG**).

Games ending with 12 point differences or less follow approximately uniform distributions. This is a notable fact: a typical final score does not exist, with roughly the same probabilities between 1 and 11 points. We define these kinds of games as **close games** (**CG**).

For a final difference higher than 12 points there is superiority of one team over another and these games are more predictable than those above. As mentioned, we define these types of game as **non-close games** (**NCG**).

In the sample analyzed a total of 3468 games (63%) concluded with a difference between 1 and 11 points and 2285 games (37%) had a difference higher than 12 points. 380 games were tied.

3. A mixture of two distributions

In order to model the number of events in the entire sample in a given time period ΔT, we assume that they follow a Poisson process, and therefore we are able to calculate the mean and variance of the number of events obtained over time unit (ΔT) for all the games. The number of seasons analyzed gives us more than 6000 games, therefore a sample large enough to measure small time intervals ΔT. The ID changes as we can observe for different ΔT (Table 1), indicating that there are several ways to observe every game. Now we analyze in depth the most clear Poissonian case $\Delta T = 10$ s, $ID = 0.98$.

Fig. 1 shows the semilogplot of the histogram of baskets scored every 10 s for both teams in the entire sample (○) and for the close games (+).

The frequency values in the histogram follow, in general, a Poisson distribution $\lambda = 0.386 \pm 0.001$ (dashed line (•)). The first five values seem to fit the Poisson distribution well. For more than five points the Poisson distribution does not fit the data well, nevertheless from two points scored the data fit better by a Power Law (solid line) with $\alpha = -6.54 \pm 0.62$; $R^2 = 0.99$, *p value* < 0.005. The α value is based on less than a data decade. It shows a fast decay of the number of events that can

Table 1
ΔT and mean values of ID.

ΔT (s)	*ID*
2	1.23
6	1.10
10	0.98
30	0.83
60	0.84

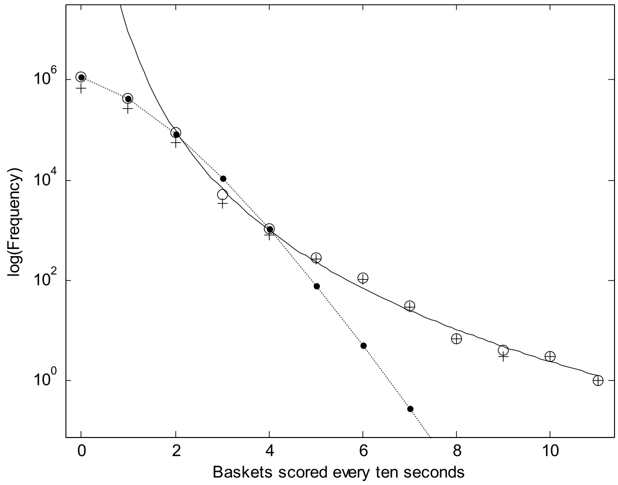

Fig. 1. Semilogplot of the histogram of all events (○) and the close games events (+) every 10 s. The dash line (●) represents the theoretical Poisson distribution for these values and the solid line shows the Power Law fit. The scaling analysis reveals that beyond a specific value the data are better fitted by a Power Law (solid line).

be obtained in this time interval. Even so, for longer intervals it remains analogous behavior. The Power Law only reveals a scale free behavior beyond a certain score value. The data appear to be distributed as a mixture of two distributions which overlap each other between events 2 and 4.

In a deeper analysis, we can observe that from a total of 1 765 728 events (including zeros), in 99.91% of cases, a maximum of 4 baskets was scored every 10 s; this is the area where the distribution matches with the Poisson distribution. Beyond this point the distribution gives the impression that it follows a scale-free behavior. At first sight, scoring more than 4 baskets in 10 s can be considered as an irrelevant event (less than 0.09% of cases). However, when we study these data in detail, we find that there are 1586 cases of 3 baskets or higher every ten seconds, where in 969 of these cases (61%) these events take place in the last minute. In addition 283 (17%) are distributed within last minutes of each quarter. Only on 333 occasions (21%) do these events come to pass in other instants of the game.

Furthermore, of the total of 1381 games with at least one case with more than 3 baskets, 1128 (82%) were in CG. As we can observe in the box of Fig. 1, in the tail of the distribution of the contested games (○) fits with the whole values (·).

We can therefore deduce that the number of baskets scored in a given time interval can be described by a mixture of two distributions. A few events in the last minute of the games seem to be the determining factor in the final results. In CG, the last minute is crucial: teams must maintain a high level of performance to reach the final minute where nothing is decided; a completely different game takes place in these last moments. A minute can be much longer than 60 s.

3.1. Scoring time evolution

To better understand the limitations of the Poisson Model, we assume that every minute ($\Delta T = 60$ s, $ID = 0.84$ *according our sample*) in a basketball game evolves following a Poisson process. Consequently, we could calculate the probability of the number of events in the first minute for all of the games, then in the second minute, and the third and so on. Every minute is considered an independent Poisson process with defined mean and variance values, and consequently ID values. Fig. 2 displays the Index of Dispersion for each minute, with values lower than 1 in most cases, and hence more predictable than the pure Poisson process. The *ID* increases with the evolution of the game, indicating more randomness. In particular, over the last minutes of the three first quarters the *ID* converges towards 1; as well minute 47, which has the value 1. The upper panels illustrate the frequency of the number of events in a semi-log plot (dotted solid line) for two cases: minute 20 and minute 47. We compared both to the Poisson distribution (dashed line (○)). In the first case the variance is lower than that predicted by Poisson. Minute 47 behaves similarly to the final minutes in previous quarters and matches the Poisson distribution, but for some tail values with low weights in the total distribution. Minute 47 represents the most

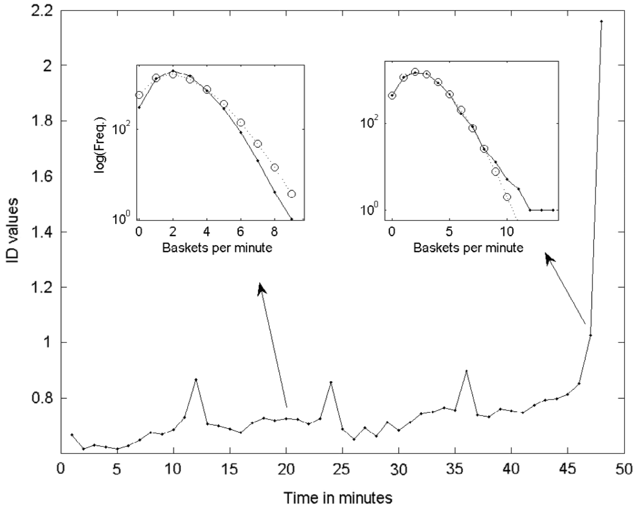

Fig. 2. Index of Dispersion of the events per minute. We observe that the trend of the values is to rise over time. Only at the end of each quarter there is a significant increase towards 1. At minute 47 the value 1 is reached (pure Poisson). Minute 48 is completely out of the range of the rest of the game, reaching values higher than 1. The behavior for this minute is very complex. The two upper panels show baskets frequency semi-log plots (dotted solid line) and Poisson distribution (dash line (○)) for minute 20 (left) and minute 47 (right).

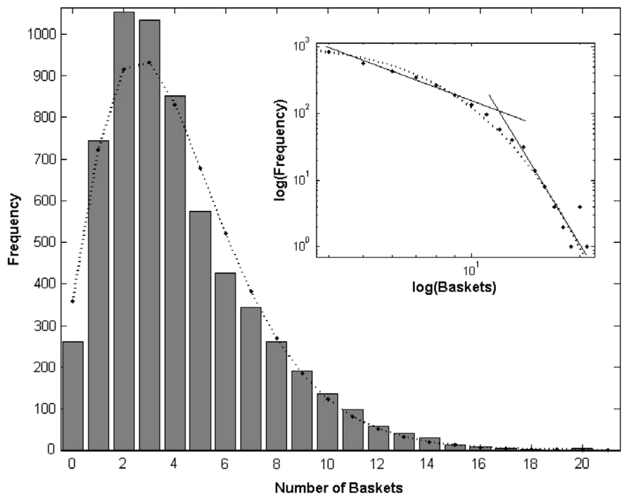

Fig. 3. Histogram of events in the last minute of the game. The solid line represents a Negative Binomial Distribution fit. Note that the values are better fitted at the tail of the distribution. For further analysis, we show a log–log plot in the upper panel, which displays two Power Laws (straight lines). The dashed line in the upper panel represents the negative binomial fit.

unpredictable moment of the game and seems to be a transition to minute 48, which is totally different from the rest of the game and will be examined in depth below.

This seems to be an indicator of the risk assumed by all teams. Teams tend to risk more during final moments of each quarter, and defensive intensity increases (more fouls) which indicates greater likelihood of failures, more zeros than at previous times or greater numbers of baskets (longer tails), meaning more randomness that explain its proximity to the Poisson distribution. In the cases with less risk, the game seems more predictable and the number of failures is lower; which would justify the least number of zeros and the shorter tail.

3.2. The last minute

We bear in mind that there is a time limit of 60 s, from 8 events the PL is truncated. Even it is possible reaching 20 events or more in the last minute, which points out the complexity of the game and importance of the last minute. Although we strictly cannot mention a long-tail phenomenon, we can do mention that there is a scale-free process.

In the last minute the variance is significantly higher than the mean, indicating an over-dispersion which can be modeled by a Negative Binomial Distribution. Fig. 3 shows a histogram of baskets and the Negative Binomial fit (parameters $r =$

Table 2
Events in the last minute.

Range of events	Points scored (%)		
	1 pts.	2 pts.	3 pts.
0–3 events	35.6	49.3	15.1
4–9 events	63.6	27.3	9.1
More than 10 events	73.1	18.8	8.1
Minute 48 total events	58.3	31.4	10.3
Total events in game time	34.0	55.0	11.0

3.86 ± 0.27; $p = 0.48 \pm 0.02$) which visually give us the impression of an acceptable model fit. The source of this behavior could be the clustering of events focused within some specific games where the numbers of basket increase. We carried out a scaling analysis on a log–log plot (upper panel) which reveals that the data can be fitted by two PLs as well. Noting that we keep in mind that there is a time limit of 60 s, from 8 events the resulting PL is truncated. Moreover, it its possible to find cases where 20 baskets or more are scored in the last minute, which points out the complexity of the game and importance of the last minute. Although we cannot strictly mention a long-tail phenomenon, we can mention that there is a scale-free process.

In order to better understand the complexity, we can distinguish the events according to the kind of point scored. In Table 2, we break down the events corresponding to the kind of point (1, 2, and 3) because until now we have only considered events and not the kind of point.

In order to better understand the game during the last minute, we can distinguish the events according to the kind of point scored because until now we have only considered events and not the kind of point.

In Table 2, we show the number of events achieved by both teams in the last minute of the game, in comparison with the game total (last row). The first column represents the number of events in ranges 0–3; 4–9; more than 10 events; the final minute and the game total. Columns 2 through 4 display the relative frequencies of points (1, 2 and 3 points) in percentage.

Note that for between zero and three events (row 1 in Table 2), the distributions for one, two or three point shots are similar to any other stage of the game (last row), except that the percentage of triples is a little bit higher. Between 4 and 9 (row 2), which roughly corresponds with the first straight line of the upper plot of Fig. 3, the number of free throws is higher. As a final point, in games with more than ten baskets in the last minute, (second straight line in Fig. 3), 73% were free throws. These two cases (rows 2 and 3) are cases of great complexity, dominated by fouls and remaining time as a stress factor. This increase in complexity (fouls committed) could be responsible for the PL behavior. For more than 9 baskets it is likely that the last minute has almost been consumed and it is more difficult to increase the number of points. This indicates truncation of the PL, as shown in Fig. 2. The Power Law indicates a longer Poisson tail, a scale-free situation, more competitiveness, intentionality, strategy and hence more rich behavior.

As Andriani and McKelvey [14,15] suggest, within competitive systems Power Law type distributions can show up under two conditions: when competitive intensity increases and when the cost of connections or information transmission decreases, which in our case would be given by better self-organization or understanding among team members over the last minute of CG.

4. The distribution of time intervals between events revisited

The second feature of the Poisson model pointed out above refers to the time interval distribution ($\mathrm{d}t$) between events. Some papers [7,10,12] show the distribution of time intervals between events ($\mathrm{d}t$). In Fig. 3, we show the same distribution on a semi-log-plot, but with an important difference. In the upper panel we highlight the details for $\mathrm{d}t < 30$ s. Note that beyond 26 s the data follow an exponential distribution (solid line in the figure and μ value) with $\mu = 0.048$. We note [10] that after 100 s, the probability of remaining without scoring is greater than that provided by the exponential distribution, this seems to be a memory effect that the exponential distribution does not have (*memorylessness*). In other words, if the game becomes complicated and neither team scores after a critical time, the probability of not scoring in the next few seconds seems to increase, which occurs in the first minutes of each quarter and especially in the fourth quarter.

Now, we focus on the area with event time differences ($\mathrm{d}t$) lower than 30 s; subplot in Fig. 4. The distribution presents three well differentiated areas marked with circles. The maximum of the distribution is located around $\mathrm{d}t = 18$ s, which seems to be related to the possession time in basketball [7,10]. The frequency values which are placed around $\mathrm{d}t = 6$ s show a change in the growth of the distribution. Also, it is especially noteworthy that 1 s differences between events (first circle) have a significant presence in the distribution. It does not seem logical that in a sport where there is certain continuity in the game, baskets with 1 s difference between them, are so numerous. However, this occurs mainly over the last minute and corresponds mostly to fouls (free throws), giving the last minute a different reality, as discussed previously.

In Fig. 5 we show a histogram of events for each $\mathrm{d}t$ value ($\mathrm{d}t < 50$ s). The dotted line ($\cdot$) corresponds uniquely to the relative frequencies of minute 48. The solid line represents the frequencies of the final minute of each quarter and minute 47 (minutes 12, 24, 36 and 47). The dashed line displays the rest of the minutes. In the upper panel the relative values of frequencies for the CG (dotted line) and NCG (solid line) are shown.

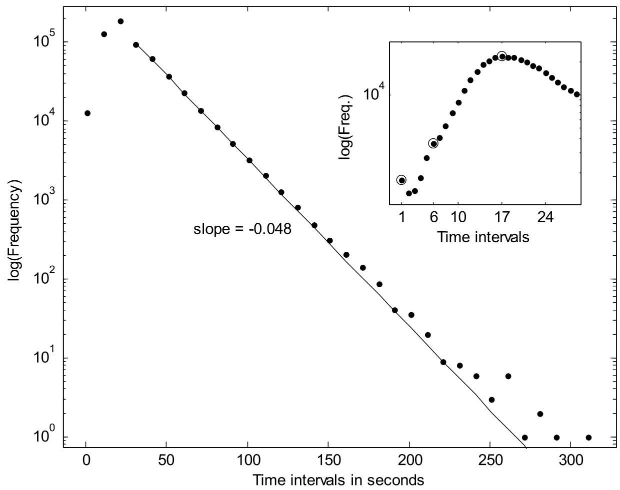

Fig. 4. Semi-log plot of time interval frequency (dt). Apparently, there are three different behaviors. The data follow a distribution with a maximum (peak) around 20 s. Beyond 30 s there is an exponential distribution. The upper panel shows a detail for the first 30 s.

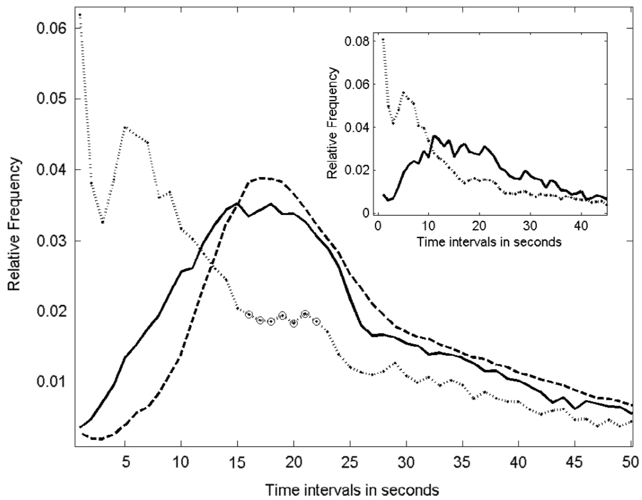

Fig. 5. Relative frequency of events (y-axis) with a time difference between them of 1 s, 2 s, 3 s, ..., up 50 s (x-axis). The dotted line (.) represents events in the last minute of the game, where it is clear that most baskets with one second difference are given in this period. The other values correspond to the last minute of the remaining quarters and minute 47 (solid line) and to the rest of the minutes (dashed line). We also highlight the area from 16 to 23 s (∘), this array has almost the same probability among its members. The relative values of frequencies for the CG (dotted line) and NCG (solid line) are displayed in the subplot.

The last minute of the game (dotted line) is completely different from the rest of the last minutes of the previous quarters. The highest frequency value is for points with dt = 1 s with a significant difference compared to the other quarters. The frequency falls by down to 3 s, but increases again to 5 s and is still high compared to the rest of sample, which coincides with the discontinuity shown in the panel of Fig. 3. There is also a nearly horizontal region at $16 < dt < 22$ that corresponds to the time differences associated with the 24 s of possession. This could be caused by the influence of the 24 s rule. Beyond this area, the frequency declines almost exponentially until the end.

Regarding the other two curves, their profiles are similar to the general one shown in Fig. 3. It is remarkable that in the case of the final minute of each quarter and in the range from 14 to 23 s, the distribution is fairly uniform, indicating a more intensive use of the 24 s rule, obeying game strategies that determine the temporal profile that we will see later.

The upper panel of Fig. 5 shows the relative frequencies of the time intervals in the last minute for the CGs (dotted line) and NCGs (solid line). Clearly the CGs are different from NCGs, which present a similar profile to the rest (see main figure). The CG curve behaves in a similar manner to the last minute curve, decreasing exponentially as a function of dt, but not homogeneously. In the interval $16 < dt < 22$, remarkably the distribution is almost uniform between these two points, in response to the kind of game performed in these situations; a shooting retention type in order to consume more time, presumably.

Table 3
Percentage of each event and points scored for different dt.

Minute of the game	dt = 1	dt = 2	dt = 3	dt = 4	dt = 5	dt = 6	dt = 7	dt = 8
11–12	1.52	4.38	5.45	6.28	5.38	5.13	5.12	4.26
23–24	3.45	5.99	7.92	7.98	7.92	6.44	6.27	5.56
35–36	2.98	3.76	5.96	6.50	6.16	4.15	4.88	4.49
46–47	3.39	4.61	6.40	6.61	6.27	5.48	5.89	5.30
47–48	69.68	57.07	45.64	40.53	32.36	22.80	19.96	12.95
Remaining game time	18.98	24.19	28.63	32.10	41.91	56.00	57.89	67.45
1 point	89.18	69.39	54.76	45.54	44.78	39.97	39.62	37.49
2 points	8.01	20.84	33.09	42.13	44.26	48.98	48.65	50.27
3 points	2.81	9.08	12.15	12.33	10.96	11.06	11.73	12.24

Table 4
Percentages of the type of points scored for a range of dt values, in CGs.

Points made	dt (s)				
	0–4	4–6	7–15	16–23	24–60
1 point	94.02	43.73	47.55	39.77	38.56
2 points	3.77	39.68	38.05	47.73	47.02
3 points	2.19	16.58	14.38	12.48	14.40

4.1. Relevance of game time and scoring

We have analyzed the behavior of time differences between events, but it will be interesting to also analyze in detail, when these time differences take place, and what kinds of points are associated with each time difference, especially in the last minute.

Table 3 shows the percentage of events with the time difference between them, dt = 1, 2, . . . , 8 s; and the minute in which they take place (column 1). Rows from 1 to 5 refer to specific minutes of the game. "Remaining game time" represents the values for remaining minutes of the game. The last three rows display the result for the same time differences for each type of point (1, 2 or 3).

Most event types with dt from 1 to 3 s take place in the last minute of the game. On the other hand, during the rest of the game (Remaining Game Time row) the behavior is the opposite so that from dt = 4 the percentage increases.

Regarding the kind of point in each situation, from dt = 1 to 3 free throws are the main type. From dt = 5 the trend is reversed, the percentage of 2 point shots starts to be greater than one point shots. This is reflected in the change of growth around the central circle in the subplot of Fig. 4.

Table 4 shows the percentages of the types of points (1, 2 or 3) for selected dt time intervals only in CGs in the last minute.

For time intervals between events dt < 4 s, in minute 48, most points are free throws. This means that teams use fouls in order to stop the game with the aim of preserving time in order to score and reduce or maintain advantage. For 4 < dt < 6, points are more balanced between free throws and two point shots. These time ranges could match with fastbreaks, where players score by a dunk or lay-up, or even a three point basket, probably following coach instructions. The aim is to have enough time to reduce the score advantage, in the case of the defending team, or take the lead from a rival failure and increase the advantage, in the case of the attacking team. Obviously fouls take place, but in a fast break a foul seems less likely than in other situations, although this is relative to score and remaining game time. For 7 < dt < 15 the case is very similar to the previous one. The range 16 < dt < 23 is the one most associated with the 24 s possession rule. In this case it seems that teams tend to retain the ball in order to ensure a good shot. But ball retention is not only an attacking strategy; it is also a defensive strategy. Retaining the ball enables the team to prevent the rival team scoring. This is a remarkable fact that can be the source of the homogeneous distribution for the range 16 < dt < 23 (figure with (◦)). We observe that the percentage of two point shots is higher than any other; the three point shot percentage decreases from the previous time range. Both situations indicate pre-established plays or coach instructions. The free throws correspond, probably, with attempts to stop the time and dominate ball possession. We have to bear in mind that retention of the ball by the attacking team can cause an increase in defensive intensity and probability of committing a foul by the defensive team.

As a final point, Fig. 5 indicates that the last minute of an NCG behaves the same as any other end of a quarter and, even, like the rest of the game.

5. Conclusions

The Poisson distribution explains most points scored in any game, in most game situations. But in CG, especially at critical times such as the last minute, the number of points is distributed following a PL, possibly an avalanche phenomenon; there are only a few such cases but they are adequately distributed among games and minute 48. Consequently the number of basket scored in a defined time interval in basketball, seems to follow a mixture of two distributions.

Therefore, the final score of close basketball games appears to depend on the last minute of the game. Furthermore, in the final minute most of these events are associated with fouls such that the likelihood of obtaining time differences of a second between them is high. The kinds of points seem to play a different role in the game. This is a characteristic of the basketball game, where points obtained from fouls play a completely different role. In fact, from the complexity point of view, free throws are emergent phenomena in a basketball game, where strategies or voluntary acts are involved with the aim simply to stop the clock.

Furthermore, the 24-second rule dynamic changes at different times of the game, especially in the last minute of each quarter and the final minute; which shows the effect the rule has on a competitive complex system, and how the elements of the system have to deal with this to obtain an advantage. This time restriction increases complexity. We verify that indeed rules do play an important role in non-linear complex systems. These changes affect game strategies, game dynamics, team dynamics, training designs, etc. In fact this is widely studied in the Exercise and Sport Sciences.

Our results show that a basketball game can be considered a quasi-pure Poisson process, not only in close games where, in general, the variance is somewhat lower than the mean, but also in non-close games. These kinds of games become more predictable as the game progresses, due to the domain that one team is exerting on the other, even though the time intervals between events and the types of scored points are completely random processes. Analyzing the reasons why one team becomes dominant over another as game progresses is very interesting to the managers of this sport.

But in games of great rivalry, the matter is quite different. Teams have to work hard to keep within a reasonable distance of the rival and, finally, confront the last minute: a completely different game. The behavior at this moment does not seem to follow a Poisson process. The emergence of scaling laws seems to suggest that the system has some kind of underlying dynamics that are different from the rest of the game, with possible avalanche phenomena and the emergence of different game situations. In addition, unlike non-close games, the outcomes of these games at these moments are unpredictable. Rivalry increases because it is still possible to win even though the score may not be favorable. Hence a well-coordinated and effective defensive strategy rises up against the desperate attempts to break the defense of the team which is losing.

It is also noteworthy how some concepts or metaphors from complex systems seem to be involved in this system, such as the Red Queen Hypothesis. But other models related to non-linear complex systems, such as "sandpile" models, in Self-Organized Critically, could also be analyzed in the case of NBA basketball. Thus, a further study might help to better understand these concepts and, in turn better, understand the problems of team sports.

The assertion attributed to Aristotle: "The whole is greater than the sum of its parts" could be translated in basketball to "minute 48 is more than the sum of 60 s", as all sport fans know and suffer.

References

[1] S. Chatterjee, M.R. Yilmaz, The NBA as an evolving multivariate system, Amer. Statist. 53 (1999) 257–262. http://dx.doi.org/10.2307/2686106.

[2] Y. de Saá Guerra, J.M. Martín González, S. Sarmiento Montesdeoca, D. Rodríguez Ruiz, A. García-Rodríguez, Juan Manuel García-Manso, A model for competitiveness level analysis in sports competitions: Application to basketball, Physica A 391 (2012) 2997–3004. http://dx.doi.org/10.1016/j.physa.2012.01.014.

[3] S.J. Gould, The streak of streaks, Chance 2 (1989) 10–16. http://dx.doi.org/10.1080/09332480.1989.10554932.

[4] T. McGarry, D.I. Anderson, S.A. Wallace, M.D. Hughes, I.M. Franks, Sport competition as a dynamical self-organizing system, J. Sports Sci. 20 (2002) 771–781. http://dx.doi.org/10.1080/026404102320675620.

[5] H.V. Ribeiro, R.S. Mendes, L.C. Malacarne, S.P. , P.A. Santoro, Dynamics of tournaments: the soccer case—A random walk approach modeling soccer leagues, Eur. Phys. J. B 75 (2010) 8. http://dx.doi.org/10.1140/epjb/e2010-00115-5.

[6] M.R. Yilmaz, S. Chatterjee, Patterns of NBA team performance from 1950 to 1998, J. Appl. Stat. 27 (2000) 555–566. http://dx.doi.org/10.1080/02664760050076399.

[7] A. Gabel, S. Redner, Random walk picture of basketball scoring, J. Quant. Anal. Sports 8 (2012). http://dx.doi.org/10.1515/1559-0410.1416.

[8] J. Greenhough, P.C. Birch, S.C. Chapman, G. Rowlands, Football goal distributions and extremal statistics, Physica A 316 (2002) 615–624. http://dx.doi.org/10.1016/S0378-4371(02)01030-0.

[9] S. Picoli Jr., R.S. Mendes, L.C. Malacarne, q-exponential, Weibull, and q-Weibull distributions: an empirical analysis, Physica A 324 (2003) 678–688. http://dx.doi.org/10.1016/S0378-4371(03)00071-2.

[10] Y. de Saá Guerra, J.M. Martín Gonzalez, S.S. Montesdeoca, D. Rodriguez Ruiz, N. Arjonilla López, J.M. García Manso, Basketball scoring in NBA games: An example of complexity, J. Syst. Sci. Complex. 26 (2013) 94–103. http://dx.doi.org/10.1007/s11424-013-2282-3.

[11] A. Clauset, M. Kogan, S. Redner, Safe leads and lead changes in competitive team sports, Phys. Rev. E 91 (2015) 62815. http://dx.doi.org/10.1103/PhysRevE.91.062815.

[12] S. Merritt, A. Clauset, Scoring dynamics across professional team sports: tempo, balance and predictability, EPJ Data Sci. 3 (2014) 4. http://dx.doi.org/10.1140/epjds29.

[13] B. Skinner, The price of anarchy in basketball, J. Quant. Anal. Sports 6 (2010). http://dx.doi.org/10.2202/1559-0410.1217.

[14] P. Andriani, B. McKelvey, Beyond Gaussian averages: redirecting international business and management research toward extreme events and power laws, J. Int. Bus. Stud. 38 (2007) 1212–1230. http://dx.doi.org/10.1057/palgrave.jibs.8400324.

[15] P. Andriani, B. McKelvey, Perspective—From Gaussian to paretian thinking: Causes and implications of power laws in organizations, Organ. Sci. 20 (2009) 1053–1071. http://dx.doi.org/10.1287/orsc.1090.0481.

[16] E.P. White, B.J. Enquist, J.L. Green, On estimating the exponent of power-law frequency distributions, Ecology 89 (2008) 905–912. http://dx.doi.org/10.1890/07-1288.1.

[17] D. Sornette, Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization, and Disorder: Concepts and Tools, Springer, 2004.

[18] M. Newman, Power laws, Pareto distributions and Zipf's law, Contemp. Phys. 46 (2005) 323–351. http://dx.doi.org/10.1080/00107510500052444.

[19] P. Bak, How Nature Works: The Science of Self-Organized Criticality, second ed., Copernicus. Springer-Verlag, New York, 1999.

[20] M.P.H. Stumpf, M.A. Porter, Critical truths about power laws, Science 335 (2012) 665–666. http://dx.doi.org/10.1126/science.1216142.

[21] A. Clauset, C.R. Shalizi, M.E.J. Newman, Power-law distributions in empirical data, SIAM Rev. 51 (2009) 661–703. http://dx.doi.org/10.1137/070710111.

[22] Y. Virkar, A. Clauset, Power-law distributions in binned empirical data, 2012. arXiv:1208.3524http://arxiv.org/abs/1208.3524 (accessed 09.05.13).