Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 369 (2006) 877–883
PHYSICA A
www.elsevier.com/locate/physa

Conditions for the emergence of scaling in the inter-event time of uncorrelated and seasonal systems

César A. Hidalgo R.[a,b,*]

[a]Center for Complex Network Research and Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA
[b]Helen Kellogg Institute, Notre Dame, IN 46556, USA

Received 12 September 2005; received in revised form 24 October 2005
Available online 25 January 2006

Abstract

Inter-event times have been studied across various disciplines in search for correlations. In this paper, we show analytical and numerical evidence that at the population level a power-law can be obtained by assuming Poissonian agents with different characteristic times, and at the individual level by assuming Poissonian agents that change the rates at which they perform an event in a random or deterministic fashion. The range in which we expect to see this behavior and the possible deviations from it are studied by considering the shape of the rate distribution.

Keywords: Power-law; Scaling; Inter-event time; Waiting time

Power-law scaling is often considered a sign of complexity. The independence of scale exhibited in this type of systems has fascinated many scientists who have attempted to explain the dynamics and correlations giving rise to this statistical property in different systems, such as complex networks [1–3], fractals [4] and economic fluctuations [5] (for a complete review of power-laws in nature and possible mechanisms that produce them we encourage the reader to take a look at Ref. [6]). This type of function also appears in allometric laws of ecology [7,8] and in the distribution of inter-event times of several different systems. In this last example, power-law scaling has been found in the stock exchange [9,10], earthquakes [11,12] e-mail login times [13], print job submissions [14], e-mail replies [15] and browsing patterns [16]. In all of these systems, we observe scaling, and the distribution of inter-event times goes as $\tau^{-\alpha}$ though the exponents tend to differ. Some of these exponents tend to be close to $\alpha = -1$, while another class tends to be around $\alpha = -2$. The first class belongs to systems governed by human decisions. Here it has been proposed, as a very likely candidate, a model based on priority queues, which captures this precise exponent [15,17], whereas the second class of behavior has been observed in earthquakes and the stock exchange [9–12].

*Corresponding author at: Center for Complex Network Research and Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA.
E-mail address: chidalgo@nd.edu.

0378-4371/$ - see front matter
doi:10.1016/j.physa.2005.12.035

In the past, exponentials have been used to explain power-laws. One of the mechanisms used is to consider a variable that has an exponential distribution

$$f(y) \sim e^{ay} \tag{1}$$

and look for a variable that is related to the first one through an exponential, such as

$$x \sim e^{by}. \tag{2}$$

When we ask for the distribution of x we have that it is given by

$$f(x) = f(y)\frac{dy}{dx} = \frac{1}{bx} e^{(a/b)\log(x)} = \frac{x^{a/b-1}}{b}. \tag{3}$$

This argument was first introduced by Miller [18,19] to explain the power-law of the distribution of words in texts. Here we study another simple way to extract a power-law, in this case from a fluctuating or time-dependent Poisson process. In the latter part of this work we show that seasonal behavior can also conduct to power-laws in the distribution of inter-event times. Seasonality is a phenomenon that becomes manifest in a variety of systems and the change of behavior induced by it can be enough to abandon the Poisson paradigm.

In our case we are concerned with a particular example, the distribution of inter-event, or waiting times. Here, we argue that the $\alpha = -2$ exponent should be expected when we consider the distribution of inter-event times in a population made of several regular components which are individually different, or when we consider individual components of heterogeneous behavior. For the first case, we consider that the rate that an agent performs an event in a certain time interval is given by p and the population of agents is such that we can define $f(p)$ as the distribution of agents with particular p's.

For the sake of clarity we assume to have a population of fireflies which fire at a rate p.[1] We also assume that we have enough flies to define a distribution of rates given by $f(p)$. We then proceed to ask each one of them how much time it had to wait between two consecutive flashes, and finally we make the histogram of this poll. The simplest case is the one in which all fireflies are equal and $f(p) = \delta(p - po)$, where δ is Dirac's delta distribution. In this case, the global behavior matches the individual one. Thus the inter-event time decays exponentially with a mean given by $1/p$.

If we consider fireflies that have a stable individual behavior, but as a population, have a broad distribution of rates, we would be in a situation in which the individual behavior does not match the global one. Individually, the fireflies will fire in a Poisson type fashion and we can approximate their inter-event times by their personal average, which is well-defined and representative at the individual level. Globally, we need to find the distribution of inter-event times. We can do this by simply calculating the fraction of fireflies that goes off in a time less than τ:

$$P(T < \tau) = P(1/p < \tau) = P(p > 1/\tau) = \int_{1/\tau}^{\infty} f(x)\, dx, \tag{4}$$

and then differentiating to get the probability density

$$P(T < \tau) = F(\infty) - F(1/\tau) \rightarrow P(T = \tau) = -\frac{dF(1/\tau)}{d\tau} = f(1/\tau)\frac{1}{\tau^2}, \tag{5}$$

which scales as $1/\tau^2$ and has an envelope given by the original function evaluated at $1/\tau$.

Numerically, we can simulate this situation by considering a distribution $f(p)$ and a big enough population. We can do this by picking up a particular firefly with a certain p from the distribution $f(p)$ and simulate the process until it goes off by asking at each time step if the firefly is going to fire or not. Fig. 1 shows our prediction for three different rate distributions. In the case of a uniform distribution (Fig. 1a), we have a system that behaves clearly as a power-law and has no envelope. We have also performed simulations with an increasing number of components to show that finite size scaling defines a clear region in which this behavior is

[1] In this work, we use the firefly analogy as a visual aid for the discussion. It is not the intention of the author to deal with synchronization or other type of phenomenon found in real fireflies.

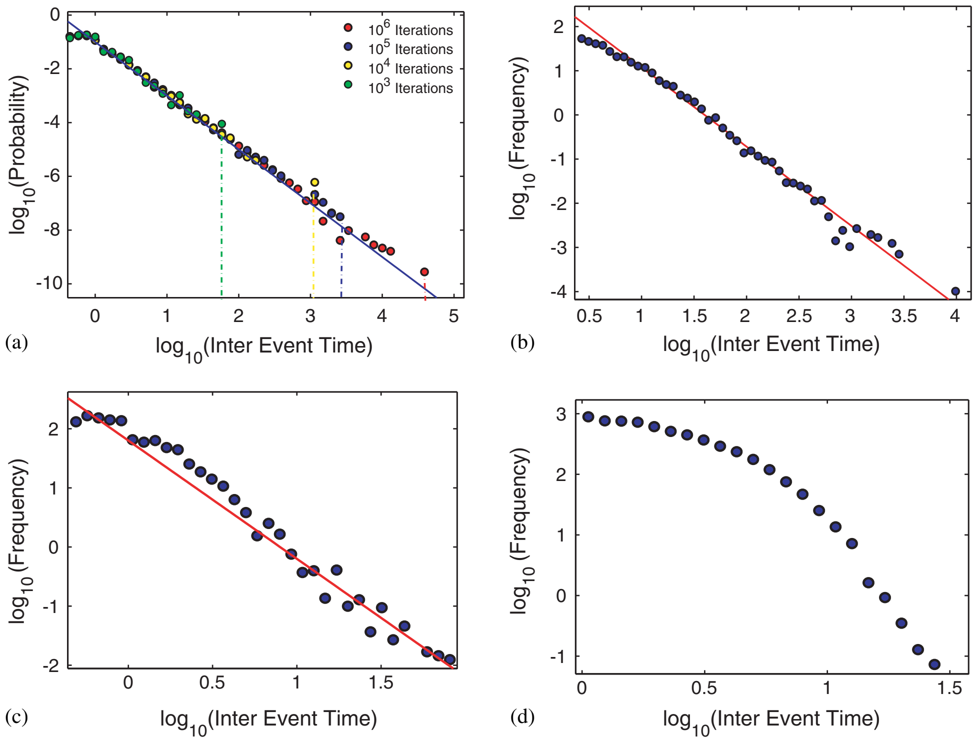

Fig. 1. (a) Finite-size scaling for the distribution of inter-event times obtained when the individual probabilities of agents were taken from a uniform distribution. The straight line has slope -2. The dash-dotted lines show the maximum inter-event time registered for a particular number of iterations. These and all subsequent plots were made with log-binning. (b) The same result is obtained when we consider an exponential distribution of rates given by $f(p) = (\frac{1}{8})\exp(-8p)$. (c) When a normal distribution is chosen for $f(p)$, the power-law behavior is present when it is wide ($N\sim[=0.4, =0.2]$) and (d) disappears when it is narrow ($N\sim[=0.4, =0.01]$).

present. From Eq. (5) we have that

$$f(p)\sim U[0,L] \to P(T=\tau) = 1/\tau^2, \tag{6}$$

and for an exponential distribution of rates we have

$$P(T=\tau) = \frac{e^{-a/\tau}}{\tau^2}. \tag{7}$$

In this case, when $\tau \to \infty$, $e^{-a/\tau} \to 1$ and the behavior is the same as for the uniform distribution which can be correctly recovered when $a \to 0$.

Deviations in the exponent can be found when the distribution of probabilities satisfies a power-law $f(p)\sim p^{\beta}$. Using Eq. (5) we can find that in this case

$$P(T=\tau) = \tau^{-(\beta+2)}, \tag{8}$$

which represents a deviation of the $\alpha = -2$ exponent which occurs in the case we inject a power-law to the system.

The studied cases do not introduce any cut-offs for large τ. This comes from the fact that long inter-event times come from small rates. All of the distributions presented above have support close to zero, so in principle times can be infinitely long. Cut-offs in the distribution can be introduced by restricting the support close to zero. A simple example of this is considering the case in which the support of $f(p)$ is restricted to the $[p_<, p_>]$. According to the formalism presented in Eqs. (4) and (5), this introduces a hard cut-off at

$\tau_{max} = 1/p_<$ for large τ and at $\tau_{min} = 1/p_>$ for small τ. Thus we can say that as a rule of thumb

$$\tau_{max} \sim 1/p_<, \quad \text{where } p_< = \min[\text{supp}[f(p)]], \tag{9}$$

$$\tau_{min} \sim 1/p_>, \quad \text{where } p_> = \max[\text{supp}[f(p)]]. \tag{10}$$

We can refine this argument for τ_{max} by considering that $f(p)$ decays in a smooth way as we approach the left edge of its support. Approximating $f(p)$ by a power series,

$$f(p) = \sum_{k=0}^{\infty} A_k p^k, \tag{11}$$

and using this in Eq. (5) we can conclude that for large τ the distribution of inter-event times goes as

$$P(T = \tau) = \sum_{k=0}^{\infty} \frac{(k+1)A_k}{\tau^{k+2}}. \tag{12}$$

When $\tau \to \infty$ (12) goes as $\tau^{k'+2}$ where k' is the coefficient of the lowest order non-vanishing expansion coefficient. To be more precise the k' exponent dominates the distribution when τ satisfies $\tau > ((k+1)A_k/A_{k'})^{1/k}\ \forall k$. To simplify this discussion, we can say that when $f(p)$ decays linearly towards zero $\alpha = -3$, and when $f(p)$ decays as a parabola we have $\alpha = -4$. The $\alpha = -2$ exponent for large times is an indication that $f(p)$ can be approximated by a constant close to the left edge of its support.

The analytical results presented above were contrasted with numerical simulations. Fig. 1b confirms that for an exponential distribution of probabilities the scaling behavior is still clearly visible and extends through several decades. We also considered the case of a normal distribution. In this case, the behavior appears when the distribution is wide enough (Fig. 1c) and disappears for narrow bells (Fig. 1d) which have negligible support close to zero.

So far, we have shown that we can expect a power-law for the distribution of inter-event times whenever we ask heterogeneous individuals for the time between its last two events and poll that data together. We have argued from simple statistics that the exponent should be $\alpha = -2$ when the distribution of rates is broad enough and it can be extended for several decades when $f(p)$ has support close to zero. This argument can be extended even further to include a population in which we do not only ask agents to tell us the inter-event time between their last two events, but we have the distribution of events for each one of them. We are concerned with the neutral case in which agents are not correlated in time or across the population, and therefore they individually follow exponential distributions. If we poll these data together by adding up all these distributions, we can also expect a power-law decay with a $\alpha = -2$ exponent. This can be seen clearly by adding up normalized exponentials representing the inter-event time distribution of regular individual agents

$$P(T = \tau) = \sum_{i}^{n} p_i \exp(-p_i \tau), \tag{13}$$

and assuming a uniform distribution of rates and a large enough population we get,[2]

$$P(T = \tau) = \int_0^{\infty} p \exp(-p\tau)\, \mathrm{d}p = -\frac{\mathrm{d}}{\mathrm{d}\tau} \int_0^{\infty} \exp(-\tau p)\, \mathrm{d}p, \tag{14}$$

which can be easily solved resulting in

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \frac{\exp(-\tau p)}{\tau} \bigg|_0^{\infty} = -\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{1}{\tau}\right) = \frac{1}{\tau^2}. \tag{15}$$

This argument can be easily generalized to include any distribution $f(p)$. In this case we have

$$P(T = \tau) = \int_0^{\infty} f(p) p \mathrm{e}^{-p\tau}\, \mathrm{d}p, \tag{16}$$

[2] This method was introduced in Ref. [13] as a possible explanation for the inter-event times, although only the case with a uniform distribution of probabilities was considered.

which may sound redundant given Eqs. (4) and (5). In fact, we introduce both methodologies because the first one is easier to use in some cases in which the integral given in Eq. (16) is not trivial to calculate. The conceptual difference of the two methods for calculating inter-event time is that in the first one we assume that the times coincide precisely with their expected values in the exponential distributions $(1/p)$, whereas in the second one we consider all possible values associated with a given probability. In the case that $f(p)$ is an exponential distribution normalized in the [0, 1] interval we have that

$$P(T=\tau)=\frac{1}{(a+\tau)^2}, \tag{17}$$

whereas when $f(p)$ is a power-law this second method requires us to solve

$$P(T=\tau)=\int_0^1 p^{\beta+1}\mathrm{e}^{-p\tau}\,\mathrm{d}p, \tag{18}$$

which can be expressed in terms of a Γ function as

$$P(T=\tau)=\left(\frac{1}{\tau}\right)^{\beta+2}\Gamma(\beta+2) \tag{19}$$

being hardly more useful than Eq. (8).

The arguments used so far have been used to show that at a population level we are likely to observe a power-law distribution of inter-event times when we poll up a large population of non-identical users. So far we have assumed that single users perform events at a fixed rate and act accordingly. But what can we expect if we allow individual agents to vary? Going back to our firefly analogy, we can imagine that we measure the time gap between several consecutive flashes for an individual firefly and discover that the histogram of inter-event times is a power-law with an exponent $\alpha=-2$. We can show that in this case, an uncorrelated random process can also explain the scaling exponent. For these matters, let us consider a firefly that initially flashes at a rate given by p_0. After it fires, we record the inter-event time and start waiting again. If p_0 is fixed, the inter-event time will decay exponentially, but if we allow this rate to change in time this would not be necessarily true. The simplest case would be to let p_0 evolve in a purely random fashion, in other words, after the firefly fires, we randomly draw a new rate p_1 from the $[p_<, p_>]$ interval. If this is the case, this system would be the same as the one in which we consider several agents with different probabilities, and therefore, it is again obvious to expect a $\alpha=-2$ exponent in the inter-event time distribution. Thus, an agent that varies its behavior in a random way maps the same model as a population of significantly different users. The way in which an individual agent varies its behavior is actually not important, as long as different rates are chosen. In fact, we can relax the assumption that rates vary randomly and instead choose a periodic function for them. In Fig. 2 we show an example of this process in which we simulate a system in which the rate at which an event is performed changes from 0.2 to 0.02 to 0.002 and then is reset back to its original value of 0.2 to start all over again. Here the rate depends only on time and does not change after an event is registered (Fig. 2a). The inter-event times are therefore correlated and the system changes periodically from activity to inactivity (Fig. 2b). The distribution of inter-event times mimics a power-law with $\alpha=-2$, but upon closer look, one can identify three humps which coincide with the expected times of the three fixed probabilities involved in the process. In order to mimic a power-law with seasonality one needs to consider $p(t)$ such that the values taken by this function are widely distributed[3] and that the function has regions in which it varies slowly enough (or not at all) to allow inter-event times to be consistent with the ratesproperly corresponding to each time.[4] Figs. 2a–c correspond to a case in which $p(t)$ mimics an exponential distribution; this is because after a linear time it decays an order of magnitude. We can shift the exponent in this case by working with a $p(t)$ that mimics a power-law. As an example we show the case in which we consider the same three probabilities as before (0.2, 0.02 and 0.002) but instead of lasting the same amount of time each, we make them last 100, 1000 and 10 000 time steps, respectively. In this case, the longer inter-event times are as frequent as the shorter ones and the system mimics a power-law with $\alpha=-1$ (Fig. 2f).

[3]We mean widely in a logarithmic sense. A similar number of values per order of magnitude.

[4]An exaggeration of this is to consider a two-step function with values 0.1 and 0.01 that has a period of two time steps. In this case, it is obvious that the system is going to be dominated by $p=0.1$ because the period of the function is shorter than the expected time associated with $p=0.01$.

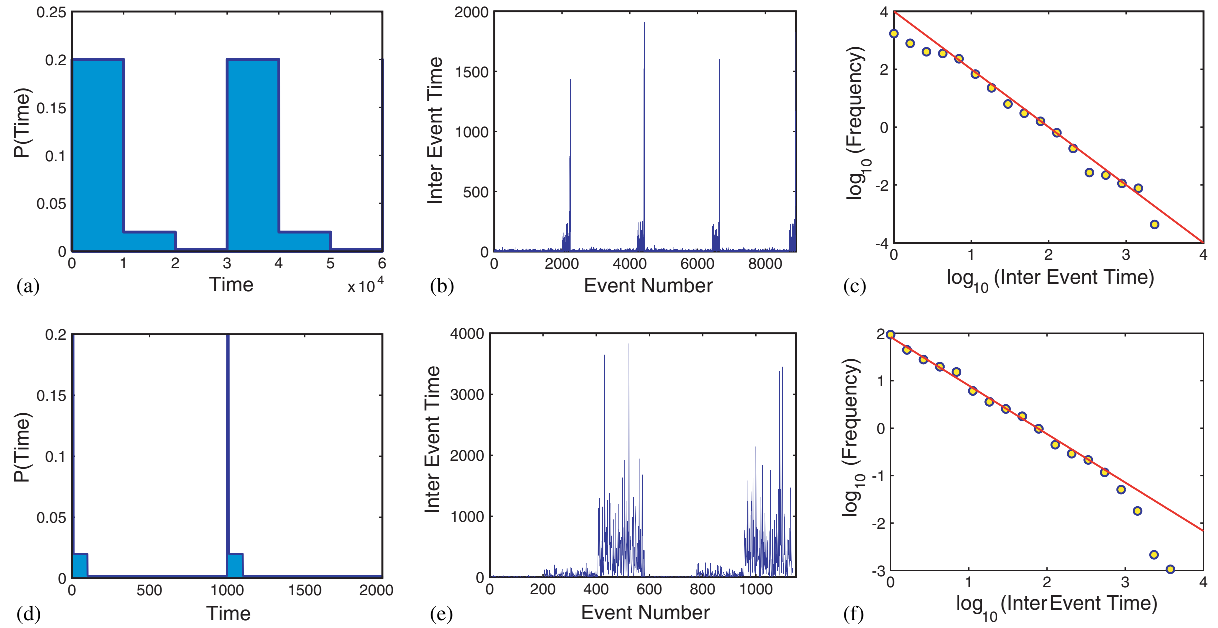

Fig. 2. (a) Periodic behavior of the rates used to model seasonal events. (b) Inter-event times obtained for a process modeled using the rates in (a). (c) The distribution of inter-event time mimics a power-law with an exponent close to -2 (straight line has slope -2). (d) The same as (a) except that in this case smaller rates are active for longer times. (e) Inter-event times. (f) For this case the distribution of inter-event times mimics a power-law with an exponent close to -1.

In the light of the previous results and examples, we have shown cases in which we can obtain power-laws by adding exponentials. The cut-off of the power-law depends on the minimum of the support of the distribution $f(p)$ and the exponent for large τ depends on the leading term of the power series expansion of $f(p)$. In the case that the function approaches the left edge of its support as a constant, we expect the exponent for large τ to be $\alpha = -2$, but in the case that the function approaches this edge as a function with singular support, we have a decrease on the exponent which is equal to the order of the singularity. Finally, when the function vanishes on the edge as a polynomial, α increases by the degree of the k' term of the polynomial for $\tau > ((k+1)A_k/A_{k'})^{1/k}\ \forall k$.

In the case of large earthquakes, it has been shown that the inter-event time of earthquakes larger than a given magnitude scales as a power-law with a $\alpha = -2$ exponent [11]. Omori's law is not valid for the inter-event time between earthquakes larger than a certain magnitude, which is the case in which you see this exponent. This is because Omori's law deals with the aftershocks which are clearly correlated. It was also argued that the $\alpha = -2$ exponent indicates a time correlated behavior, because the inter-event time distribution is not Poissonian. From the analytical arguments shown above, we can see that the $\alpha = -2$ exponent is precisely what is expected for the uncorrelated case in which we consider that an earthquake occurs with a probability that is randomly reset after each event. In the case of the stock exchange [10], it was shown that the scaling exponent tends to decrease as the threshold on the normalized fluctuations increases. In other words, when waiting times between large fluctuations are considered, the scaling exponent approaches $\alpha = -1$ indicating possibly a different mechanism or a power-law injection, whereas when small variations are considered the exponent is close to $\alpha = -2$, which could be a signature of seasonality [9] or uncorrelated probability variations.

Despite the simplicity of our calculations, Poisson process is usually assumed in stochastic modeling. Here we have shown that when we have a broad population which participates in individual Poisson process or agents which have non-stationary behavior the waiting time distributions follow a power-law. In the case of simple Poisson processes the first thing that should be considered is that usually not all of the components of the system behave in the same way; this consideration alone is enough to define a scaling region that has not usually been considered.

I would like to thank D. Menzies and the referees for helpful reviews of this manuscript.

References

[1] A.-L. Barabási, R. Albert, Science 286 (1999) 509.
[2] R. Albert, A.-L. Barabási, Rev. Mod. Phys. 74 (2002) 47.
[3] M.E.J. Newman, SIAM Rev. 45 (2003) 167.
[4] T.A. Witten Jr., L.M. Sander, Phys. Rev. Lett. 47 (1981) 1400.
[5] V. Plerou, P. Gopikrishnan, L.A.N. Amaral, X. Gabaix, H.E. Stanley, Phys. Rev. E 62 (2002) 3 (Rapid Communication R3023).
[6] M.E.J. Newman, Contemp. Phys. 46 (2005) 323–351.
[7] D. Garlaschelli, G. Caldarelli, L. Petroniero, Nature 423 (2003) 165.
[8] G. West, J.H. Brown, B.J. Enquist, Science 276 (1997) 122.
[9] E. Scalas, R. Gorenflo, H. Luckock, F. Mainardi, M. Mantelli, M. Roberto, Quant. Finance 4 (2004) 695.
[10] T. Kaizoji, M. Kaizoyi, Physica A 336 (2004) 563.
[11] M.S. Mega, P. Allegrini, P. Grigolini, V. Latora, L. Palatella, A. Rapisarda, S. Vinciguerra, Phys. Rev. Lett. 90 (2003) 18850.
[12] S. Abe, N. Suzuki, Physica A 350 (2005) 588.
[13] A. Chatterjee, cond-mat/0307533.
[14] U. Harder, M. Paczuski, Physica A 361 (2006) 329–336.
[15] A.-L. Barabási, Nature 207 (2005) 435.
[16] Z. Dezso, E. Almaas, A, Lukács, B. Rácz, I. Szakadát, A.-L. Barabási, physics/0505087.
[17] A. Vazquez, Phys. Rev. Lett. 95 (2005) 248701.
[18] G.A. Miller, Am. J. Psychol. 70 (1957) 311.
[19] W. Li, IEEE Trans. Inf. Theory 38 (1992) 1842.