Physica A 474 (2017) 312–318

Contents lists available at ScienceDirect

Physica A

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

A model-free characterization of recurrences in stationary time series

Rémy Chicheportiche [a], Anirban Chakraborti [a,b,*]

[a] Chaire de Finance Quantitative, Laboratoire de Mathématiques Appliquées aux Systèmes, École Centrale Paris, 92290 Châtenay-Malabry, France

[b] School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi-110067, India

HIGHLIGHTS

- Copula is a natural model-free framework to study non-linear dependencies like recurrences.
- Non-linear dependencies do impact both the statistics and dynamics of recurrence times.
- Scaling arguments for the unconditional distribution may not be applicable.
- Fitting and/or simulating the intertemporal distribution of recurrence intervals is very much system specific.

ARTICLE INFO

Article history:
Received 31 July 2014
Received in revised form 28 January 2016
Available online 19 January 2017

Keywords:
Recurrence intervals
Copulas
Long-ranged correlations
Time series

ABSTRACT

Study of recurrences in earthquakes, climate, financial time-series, etc. is crucial to better forecast disasters and limit their consequences. Most of the previous phenomenological studies of recurrences have involved only a long-ranged autocorrelation function, and ignored the multi-scaling properties induced by potential higher order dependencies. We argue that copulas is a natural model-free framework to study non-linear dependencies in time series and related concepts like recurrences. Consequently, we arrive at the facts that (i) non-linear dependences *do* impact both the statistics and dynamics of recurrence times, and (ii) the scaling arguments for the unconditional distribution may not be applicable. Hence, fitting and/or simulating the intertemporal distribution of recurrence intervals is very much system specific, and cannot actually benefit from universal features, in contrast to the previous claims. This has important implications in epilepsy prognosis and financial risk management applications.

1. Introduction

Extreme events are widely studied in seismicity, astronomy, physiology, finance, etc. [1]. In particular, the inter-occurrence times (or recurrence intervals), i.e. the periods between two occurrences of the observed phenomenon that exceed a given threshold, have important implications in risk management in a view to predict the advent of such extreme events or characterize aftershocks.

* Corresponding author.
E-mail addresses: remy.chicheportiche@graduates.centraliens.net (R. Chicheportiche), anirban@jnu.ac.in (A. Chakraborti).
URL: http://www.jnu.ac.in/Faculty/anirban/index.html (A. Chakraborti).

http://dx.doi.org/10.1016/j.physa.2017.01.073
0378-4371/

If we consider a time series $\{X_t\}_{t=1...T}$ of length T, as a realization of a discrete stochastic process, then the joint cumulative distribution function (CDF) of n occurrences ($1 \leq t_1 < \cdots < t_n < T$) of the process is:

$$\mathcal{F}_{t_1,...,t_n}(\mathbf{x}) = \mathbb{P}[X_{t_1} < x_{t_1}, \ldots, X_{t_n} < x_{t_n}]. \tag{1}$$

We assume that the process is *stationary* with a distribution F, and a translational-invariant joint distribution $\mathcal{F}$ with long-ranged dependences, as is typically the case e.g. for seismic and financial data.

A realization of X_t at date t is called an "event" when its value exceeds a threshold in the form of a quantile $X^{(+)} = F^{-1}(1-p_+)$, where p_+ is the small unconditional probability of crossing that threshold. As we show just below, recurrences of such events only involve the diagonal n-points probability

$$\mathcal{C}_n(p) = \mathcal{F}_{t+[\![1,n]\!]}(F^{-1}(p), \ldots, F^{-1}(p)), \tag{2}$$

that all $n \geq 1$ consecutive variables $X_{t+1}, \ldots, X_{t+n}$ are below the upper pth quantile of the stationary distribution, where $p \in [0, 1]$ and $t + [\![1, n]\!]$ is a shorthand for $\{t+1, \ldots, t+n\}$. Clearly, $\mathcal{C}_1(p) = p$ and we set by convention $\mathcal{C}_0(p) \equiv 1$. Eq. (2) in fact defines, through Sklar's theorem, the diagonal of what statisticians call the n-points "copula". A copula is nothing else than a multivariate CDF with uniform marginals, see e.g. Ref. [2].

A consensus has emerged on the statistics of recurrence times from many phenomenological studies on real or simulated data with long-ranged correlations: the unconditional probability distribution function (PDF) of recurrence intervals τ follows a scaling relation [3–6]

$$\pi(\tau) = \frac{1}{\tau_c} f\left(\frac{\tau}{\tau_c}\right), \tag{3}$$

where τ_c is a characteristic recurrence time for a given time series and a choice of event-triggering threshold. Possible and reported scaling functions include: exponential decay $\ln f(x) \sim x$ (which corresponds to independent arrivals), power-law decay $\ln f(x) \sim \ln x$, stretched exponential (generalized Gamma) $\ln \ln f(x) \sim \ln x$, and others, e.g. mixed stretched exponential + power-law, possibly due to finite size and discreteness effects [5,6]. Furthermore, the sequence of recurrence times $\{\tau_i\}$ exhibits long-term correlation $\langle \tau_i \tau_{i+\ell} \rangle \sim \ell^{-\gamma}$ [7]. Yet we show that this consensus, founded on the misconception that "nonlinearities are not needed to explain the properties studied" [4], is incorrect since the whole non-linear dependences affect $\pi(\tau)$. In particular, simple theories based on a benchmark model of triggered seismicity have also found this universality to be only approximate [8,9].

Some *ad hoc* attempts at bringing in non-linearities and/or multi-scaling have also been made, in view of modeling the behavior of a specific system, e.g. financial returns with multifractal log-volatility [10,11]. In this paper we show that copulas is a natural *model-free* theoretical framework to study non-linear dependencies in time series. This implies that non-linear correlations and multi-point dependences are relevant for the related concept of recurrences. As a consequence, a scaling relation of the form (3) is at best approximate, and would only hold for processes exhibiting a time-dependence characterized by a unique time scale τ_c. Furthermore, a characterization of clustering based on the autocorrelation of recurrence intervals is an oversimplified view of the reality.

This paper would further illustrate the use of copulas in the very recent study of serial dependences in time series and discrete processes, rather than of multivariate dependences (for what the concept was originally introduced), and complement the studies of Refs. [12–14] with an emphasis on the "no-scaling" argument, see below. It is mainly motivated by the invariance property of the copula under monotonous rescaling of the variables, i.e. the irrelevance of the marginal (stationary) distribution, for some applications we investigate here.

2. Recurrence interval studies using copulas and results

Empirically, the n-points probabilities are very hard to measure due to the large noise associated with such rare joint occurrences. However, there exist observables that embed many-points properties and are more easily measured, such as the length of sequences (clusters) of thresholded events [13], and the recurrence times of such events, that we study here using copulas.

As a first example, the Gaussian diagonal copula is

$$\mathcal{C}_n(p) = \Phi_\rho\big(\Phi^{-1}(p), \ldots, \Phi^{-1}(p)\big) \tag{4}$$

where Φ^{-1} is the univariate inverse CDF, and Φ_ρ denotes the multivariate CDF with $(n \times n)$ covariance matrix ρ, which is Tœplitz with symmetric entries

$$\rho_{tt'} \equiv \langle X_t X_{t'} \rangle = \rho(|t-t'|), \quad t, t' = 1, \ldots, n. \tag{5}$$

Although the n-points expectations of Gaussian processes reduce to all combinations of the 2-points expectations (5), their full dependence structure is *not* reducible to the bivariate distribution, unless the process is also Markovian, i.e. only in the particular case of exponential correlation. The White Noise product copula $\mathcal{C}_n(p) = p^n$ is recovered in the limit of vanishing correlations $\rho(\ell) = 0\,\forall \ell$, and other examples include the exponentially correlated Markovian Gaussian Noise, the power-law correlated (thus scale-free) Fractional Gaussian Noise, and the logarithmically correlated multifractal Gaussian Noise.

Note that copulas are invariant under any continuous monotonous transformation of the X_t's, and are thus better suited to study temporal dependences than e.g. the linear correlation function. Indeed if f and g are two increasing functions, $\langle f(X_t)f(X_{t+\tau})\rangle$ and $\langle g(X_t)g(X_{t+\tau})\rangle$ can be arbitrarily different in spite of the underlying process being the same, whereas the copulas of $\{f(X_t)\}_t$ and $\{g(X_t)\}_t$ are identical.

2.1. Intertemporal distribution of recurrence times

The probability $\pi(\tau)$ of observing a recurrence interval τ between two events is the conditional probability of observing a sequence of $\tau - 1$ "non-events" bordered by two events:

$$\pi(\tau) = \mathbb{P}[X_\tau > X^{(+)}, X_{[\![1;\tau[\![} < X^{(+)} | X_0 > X^{(+)}].$$

After a simple operation flipping all '>' signs to '<', it can be written in the language of copulas as [15]:

$$\pi(\tau) = \frac{\mathcal{C}_{\tau-1}(1-p_+) - 2\,\mathcal{C}_\tau(1-p_+) + \mathcal{C}_{\tau+1}(1-p_+)}{p_+}. \tag{6}$$

Please see Appendix for a detailed derivation of this equation.

The cumulative distribution

$$\Pi(\tau) = \sum_{n=1}^{\tau} \pi(n) = 1 - \frac{\mathcal{C}_\tau(1-p_+) - \mathcal{C}_{\tau+1}(1-p_+)}{p_+}$$

is more appropriate for empirical purposes, being less sensitive to noise.[1] These exact expressions make clear – almost straight from the definition – that (i) the distribution of recurrence times *depends only on the copula* of the underlying process and not on the stationary law, in particular its domain or its tails (this is because we take a relative definition of the threshold as a quantile); (ii) *non-linear* dependences are highly relevant in the statistics of recurrences, so that linear correlations can in the general case by no means explain alone the properties of $\pi(\tau)$ [16]; and (iii) recurrence intervals have a *long memory* revealed by the $(\tau + 1)$-points copula being involved, so that the recurrences themselves can only be memoryless in very restricted cases where this copula cancels out in the multivariate distribution (see Eq. (A.1) below).[2] Hence, when the copula is known (Eq. (4) for Gaussian processes), the distribution of recurrence times is exactly characterized by the analytical expression in Eq. (6).

The average recurrence time $\mu_\pi \equiv \langle\tau\rangle$ is found straightforwardly, and the variance $\sigma_\pi^2 \equiv \langle\tau^2\rangle - \mu_\pi^2$ of the distribution can be computed as well:

$$\mu_\pi = \frac{1}{p_+}, \tag{7}$$

$$\sigma_\pi^2 = \frac{2}{p_+}\sum_{\tau=1}^{\infty}\mathcal{C}_\tau(1-p_+) - \frac{1-p_+}{p_+^2}. \tag{8}$$

Importantly, μ_π is *universal* whatever the dependence structure.[3] Introducing the copula allows to emphasize the validity of the statement even in the presence of non-linear long-term dependences, as Eq. (7) means that the average recurrence interval is *copula*-independent. This is intuitive as, for a given threshold, the whole time series is the succession of a fixed number p_+T of recurrences whose lengths τ_i necessarily add up to the total size T, so that $\langle\tau\rangle = \sum_i \tau_i/(p_+T) = 1/p_+$. Note that Eq. (7) assumes an infinite range for the possible lags τ, which is achieved either by having an infinitely long time series, or more practically when the translational-invariant copula is periodic at the boundaries of the time series, as is typically the case for artificial data which are simulated using numerical Fourier Transform methods. The variance σ_π^2 is not universal, in contrast with the mean, and can be related to the average unconditional waiting time [13].

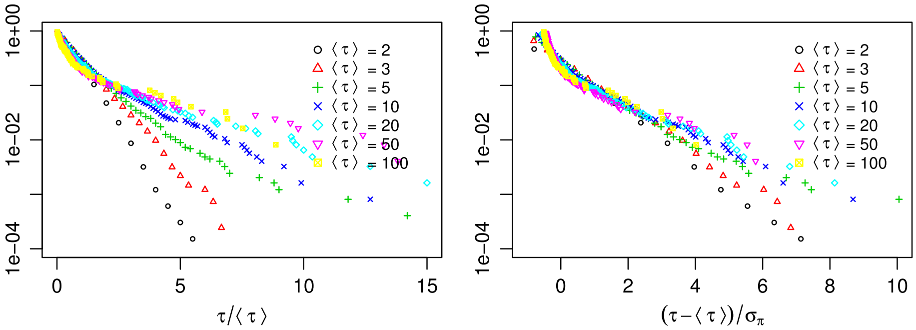
The universality of the average recurrence time has important implications for the potential scaling properties of the PDF of recurrence times. Indeed, it has been believed that $\pi(\tau)$ can be fitted by a unique scaling function Eq. (3), with f depending on the underlying process and its dependence structure. Such a scaling would be of paramount importance in the empirical investigation of extreme events (for which p_+ is close to 0 and there is often too few data points to conduct a thorough statistical study), as it would make it possible to extrapolate the distribution found at low thresholds to large ones. Now, because there is a one-to-one correspondence between $\langle\tau\rangle$ and p_+ following the universality (7), the natural scale τ_c is necessarily the mean recurrence times, and the relation (3) describes in fact a scaling of π with the threshold.

On the theoretical side, the fact that the distribution of recurrence times exhibits a universal scaling ranging over several orders of magnitude has been shown to be related to criticality in long-ranged correlated complex systems, like invariance

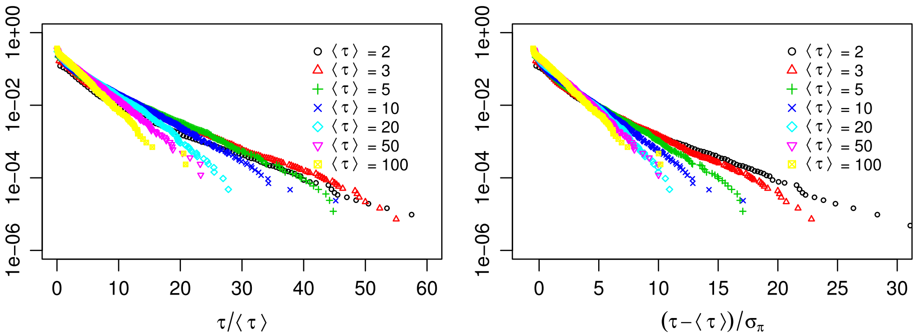
[1] If $\pi(\tau)$ has exponential tails, then $1 - \Pi(\tau)$ has the same tails; if $\pi(\tau)$ has power-law tails, then $1 - \Pi(\tau)$ has power-law tails, too. Hence representing $1 - \Pi(\tau)$ in a lin-log or log–log scale, respectively, is as meaningful as representing $\pi(\tau)$ itself.

[2] This is obviously the case for the regular Poisson process, but also for the fractional Poisson process [17] and some renewal processes [6,18].

[3] This result was first stated and proven in Ref. [19], in a similar fashion.



(a) IBM daily stock returns.



(b) EEG.

Fig. 1. (Color online) To illustrate that no scaling occurs when non-linear dependences and/or several time scales are involved in the dependence structure, we plot the tail probability $1 - \Pi(\tau)$ of the recurrence intervals versus $\tau/\langle\tau\rangle$, at several thresholds $p_+ = 1/\langle\tau\rangle$ for (a) financial data, and (b) EEG data.

under Renormalization-Group transformation [20,21]. But such a universal scaling can only exist for *linearly* correlated processes [16], and there is even a concern [22] whether there exist at all fixed-point solutions of the RG equations other than the trivial exponential function corresponding to independent arrivals. Therefore, since the average recurrence time does not carry any information whatsoever about the dependence structure of the process, a scaling relation of the form (3) is a trivial one when $\tau_c = \langle\tau\rangle$, and it might hold for processes exhibiting a time-dependence fully characterized by linear correlations, and no other relevant time scale; this is the case only for the power-law correlated Gaussian process [23] (but see [24]), for which $\pi(\tau)$ can anyways be expressed analytically according to Eqs. (4), (6) with $\rho(\ell) \sim \ell^{-\gamma}$. But no such simple scaling is expected in the general case when either non-linear dependences are present, and/or when several time scales are involved in the dependence structure.

We illustrate this on Fig. 1(a) for the daily log-returns of the IBM stock from 1962 to 2010 (same data as in Ref. [25], with however a sign flip since our definition of the threshold is upside whereas there it is downside). Financial returns have a very short term linear correlation, a very long term and multifractal correlation of amplitudes, and other kinds of intermediate-scale dependences (e.g. leverage effect) [26], that show up in the copula in a very peculiar fashion [12]. Hence, as expected [25], the distribution of recurrence times has a more complex functional form than what is allowed by the scaling in Eq. (3), as reproduced on Fig. 1(a). One may note by the way that the correlation of the amplitudes (volatility clustering) is the dominant one, so that the series of $\omega_t = \log|X_t|$ is mostly linearly correlated and in fact the recurrence intervals distribution of ω_t does exhibit an approximate scaling of the type (3) [27,28,5].

We also illustrate the non-scaling behavior on measurements of electroencephalograms (EEG) on brain surface of patients awake with eyes open (set A of the dataset studied in Ref. [29]). It consists of 100 stationary series of 4097 observations each; once centered and rescaled, all series have same distribution. The results are shown on Fig. 1(b): the presence of several (in fact periodic) scales in the dependence structure forbids the scaling of $\pi(\tau)$ with $\langle\tau\rangle$.

In passing, notice that the statistics of recurrence times is much related to that of sequence lengths: an interval τ between two events always characterizes at the same time a sequence of $\tau - 1$ "non-events". In this respect, the average sequence length can also be shown to be universal, what rules out part of the analysis of Ref. [30], where the authors use the relation (7) as a *test* of independence of the events above/below the threshold $X^{(+)} = 0$.

2.2. Conditional recurrence times

The dynamics of recurrence times is as important as their statistical properties, and in fact impacts the empirical determination of the latter.[4] It is now clear, both from empirical evidences and analytically from the discussion on Eq. (6), that recurrence intervals have a long memory. In dynamic terms, this means that their occurrences show some clustering. The natural question is then: "Conditionally on an observed recurrence time, what is the probability distribution of the next one?" This probability of observing an interval τ' immediately following an observed recurrence time τ is

$$\mathbb{P}[X_{\tau+\tau'} > X^{(+)}, X_{\tau+[\![1;\tau'[\![} < X^{(+)} | X_\tau > X^{(+)}, X_{[\![1;\tau[\![} < X^{(+)}, X_0 > X^{(+)}]. \tag{9}$$

Please see Appendix for a detailed derivation of this equation.

Again, flipping the '>' to '<' allows to decompose it as

$$\frac{\mathcal{C}_{\tau-1;\tau'-1} - \mathcal{C}_{\tau;\tau'-1} - \mathcal{C}_{\tau-1;\tau'} + \mathcal{C}_{\tau;\tau'}}{\mathcal{C}_{\tau-1} - 2\mathcal{C}_\tau + \mathcal{C}_{\tau+1}} - \frac{\pi(\tau+\tau')}{\pi(\tau)},$$

where the $(\tau+\tau')$-points probability

$$\mathcal{C}_{\tau;\tau'}(p) = \mathcal{F}_{[\![0;\tau+\tau']\!]\setminus\{\tau\}}(F^{-1}(p), \dots, F^{-1}(p))$$

shows up. Of course, this exact expression has no practical use, again because there is no hope of empirically measuring any many-points probabilities of extreme events with a meaningful signal-to-noise ratio. We rather want to stress that non-linear correlations and multi-points dependences are relevant, and that a characterization of clustering based on the autocorrelation of recurrence intervals is an oversimplified view of reality.

3. Discussions

The exact universality of the mean recurrence interval imposes a natural scale in the system. A scaling relation in the distribution of such recurrences is only possible in the absence of any other characteristic time. When such additional characteristic times are present (typically in the non-linear correlations), the rescaling must be performed with a scale-dependent renormalization, in the spirit of

$$\pi(\tau) = \frac{1}{\sigma_\pi} g\left(\frac{\tau - \langle\tau\rangle}{\sigma_\pi}\right), \tag{10}$$

where σ_π is defined in Eq. (8) and *does* contain information about the dependence structure. Then in some cases where σ_π embeds all the dependence of the underlying process, the scaling (10) might hold, at least approximately (see e.g. Fig. 1(a)), even in the presence of several scales, as the recurrence times corresponding to the different regimes would be renormalized accordingly.

We stress that recurrences are intrinsically multi-points objects related to the non-linear dependences in the underlying time-series. As such, their autocorrelation is not a reliable measure of their dynamics, for their conditional occurrence probability is much history dependent. We have reported elsewhere [13] more properties of recurrence times and the statistics of other observables (waiting times, cluster sizes, records, aftershocks) in light of their description in terms of the diagonal copula. We hope that these studies can shed light on the n-points properties of the process by assessing the statistics of simple variables rather than positing an *a priori* 2-points correlation structure and deriving a corresponding recurrence times distribution.

Acknowledgments

We thank F. Abergel, F. Baldovin, J.-P. Bouchaud, D. Challet, A. Corral, N. Millot, M. Politi, E. Scalas, G. Tilak for helpful discussions and comments. R.C. acknowledges financial support by Capital Fund Management, Paris. A.C. acknowledges the financial support by University of Potential Excellence-II grant (Project ID-47) of the Jawaharlal Nehru University, New Delhi, and grant number BT/BI/03/004/2003(C) of Government of India, Ministry of Science and Technology, Department of Biotechnology, Bioinformatics division.

Appendix

Detailed derivation of Eq. (6)

$$\pi(\tau) = \mathbb{P}[X_\tau > X^{(+)}, X_{[\![1;\tau[\![} < X^{(+)} | X_0 > X^{(+)}].$$

[4] Distribution testing for $\pi(\tau)$ involving Goodness-of-fit tests [31,32] should be discarded because those are not designed for dependent samples and rejection of the null cannot be relied upon. See Ref. [12] for an extension of GoF tests when some dependence is present.

After a simple algebraic transformation flipping all '>' signs to '<', it can be written in the language of copulas as:

$$\begin{aligned}
\pi(\tau) &= \frac{\mathbb{P}[X_\tau > X^{(+)}, X_{[\![1;\tau[\![} < X^{(+)}; X_0 > X^{(+)}]}{\mathbb{P}[X_0 > X^{(+)}]} \\
&= \frac{\mathbb{P}[X_{[\![1;\tau[\![} < X^{(+)}]}{p_+} - \frac{\mathbb{P}[X_\tau < X^{(+)}, X_{[\![1;\tau[\![} < X^{(+)}]}{p_+} \\
&\quad - \frac{\mathbb{P}[X_{[\![1;\tau[\![} < X^{(+)}; X_0 < X^{(+)}]}{p_+} + \frac{\mathbb{P}[X_\tau < X^{(+)}, X_{[\![1;\tau[\![} < X^{(+)}; X_0 < X^{(+)}]}{p_+} \\
\pi(\tau) &= \frac{\mathcal{C}_{\tau-1}(1-p_+) - 2\,\mathcal{C}_\tau(1-p_+) + \mathcal{C}_{\tau+1}(1-p_+)}{p_+}.
\end{aligned}$$

Detailed derivation of Eq. (9)

The probability of observing an interval τ' immediately following an observed recurrence time τ is

$$\mathbb{P}[X_{\tau+\tau'} > X^{(+)}, X_{\tau+[\![1;\tau'[\![} < X^{(+)} | X_\tau > X^{(+)}, X_{[\![1;\tau[\![} < X^{(+)}, X_0 > X^{(+)}]. \tag{A.1}$$

Again, flipping the '>' to '<' allows to decompose it as

$$\begin{aligned}
&\frac{\mathbb{P}[X_{\tau+k} < X^{(+)}, X_\tau > X^{(+)}, X_n < X^{(+)}, X_0 > X^{(+)}, 1 \le n < \tau, 1 \le k < \tau']}{\mathbb{P}[X_n < X^{(+)}, X_0 > X^{(+)}, 1 \le n < \tau] - \mathbb{P}[X_n < X^{(+)}, X_0 > X^{(+)}, 1 \le n \le \tau]} \\
&\quad - \frac{\mathbb{P}[X_{\tau+k} < X^{(+)}, X_\tau > X^{(+)}, X_n < X^{(+)}, X_0 > X^{(+)}, 1 \le n < \tau, 1 \le k \le \tau']}{\mathbb{P}[X_n < X^{(+)}, X_0 > X^{(+)}, 1 \le n < \tau] - \mathbb{P}[X_n < X^{(+)}, X_0 > X^{(+)}, 1 \le n \le \tau]} \\
&= \frac{\mathbb{P}[X_{\tau+k} < X^{(+)}, X_\tau > X^{(+)}, X_n < X^{(+)}, 1 \le n < \tau, 1 \le k < \tau']}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} \\
&\quad - \frac{\mathbb{P}[X_{\tau+k} < X^{(+)}, X_\tau > X^{(+)}, X_n < X^{(+)}, 0 \le n < \tau, 1 \le k < \tau']}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} \\
&\quad - \frac{\mathbb{P}[X_{\tau+k} < X^{(+)}, X_\tau > X^{(+)}, X_n < X^{(+)}, 1 \le n < \tau, 1 \le k \le \tau']}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} \\
&\quad + \frac{\mathbb{P}[X_{\tau+k} < X^{(+)}, X_\tau > X^{(+)}, X_n < X^{(+)}, 0 \le n < \tau, 1 \le k \le \tau']}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} \\
&= \frac{\mathcal{C}_{\tau-1;\tau'-1}(p) - \mathcal{C}_{\tau+\tau'-1}(p)}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} - \frac{\mathcal{C}_{\tau;\tau'-1}(p) - \mathcal{C}_{\tau+\tau'}(p)}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} \\
&\quad - \frac{\mathcal{C}_{\tau-1;\tau'}(p) - \mathcal{C}_{\tau+\tau'}(p)}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} + \frac{\mathcal{C}_{\tau;\tau'}(p) - \mathcal{C}_{\tau+\tau'+1}(p)}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} \\
&= \frac{\mathcal{C}_{\tau-1;\tau'-1} - \mathcal{C}_{\tau;\tau'-1} - \mathcal{C}_{\tau-1;\tau'} + \mathcal{C}_{\tau;\tau'}}{\mathcal{C}_{\tau-1} - 2\mathcal{C}_\tau + \mathcal{C}_{\tau+1}} - \frac{\pi(\tau+\tau')}{\pi(\tau)},
\end{aligned}$$

where

$$\mathcal{C}_{\tau;\tau'}(p) = \mathcal{F}_{[\![0;\tau+\tau']\!]\setminus\{\tau\}}(F^{-1}(p), \ldots, F^{-1}(p)).$$

References

[1] A. Bunde, J. Kropp, H.-J. Schellnhuber, The Science of Disasters: Climate Disruptions, Hheart Attacks, and Market Crashes, vol. 2, Springer Verlag, 2002.
[2] L. Rüschendorf, Copulas, Sklar's theorem, and distributional transform, in: Mathematical Risk Analysis, Springer, 2013, pp. 3–34.
[3] P. Bak, K. Christensen, L. Danon, T. Scanlon, Unified scaling law for earthquakes, Phys. Rev. Lett. 88 (17) (2002) 178501.
[4] A. Bunde, J.F. Eichner, J.W. Kantelhardt, S. Havlin, Long-term memory: A natural mechanism for the clustering of extreme events and anomalous residual times in climate records, Phys. Rev. Lett. 94 (2005) 8701.
[5] J.F. Eichner, J.W. Kantelhardt, A. Bunde, S. Havlin, Statistics of return intervals in long-term correlated records, Phys. Rev. E 75 (2007) 1128.
[6] M.S. Santhanam, H. Kantz, Return interval distribution of extreme events and long-term memory, Phys. Rev. E 78 (2008) 1113.
[7] V.N. Livina, S. Havlin, A. Bunde, Memory in the occurrence of earthquakes, Phys. Rev. Lett. 95 (2005) 8501.
[8] A. Saichev, D. Sornette, 'Universal' distribution of interearthquake times explained, Phys. Rev. Lett. 97 (7) (2006) 78501.
[9] S. Touati, M. Naylor, I.G. Main, Origin and nonuniversality of the earthquake interevent time distribution, Phys. Rev. Lett. 102 (16) (2009) 168501.
[10] M.I. Bogachev, J.F. Eichner, A. Bunde, Effect of nonlinear correlations on the statistics of return intervals in multifractal data sets, Phys. Rev. Lett. 99 (2007) 240601.
[11] A. Corral, Scaling in the timing of extreme events, Chaos Solitons Fractals 74 (2015) 99–112.
[12] R. Chicheportiche, J.-P. Bouchaud, Goodness-of-fit tests with dependent observations, J. Stat. Mech. Theory Exp. 2011 (09) (2011) P09003.
[13] R. Chicheportiche, A. Chakraborti, Copulas and time series with long-ranged dependencies, Phys. Rev. E 89 (2014) 042117.
[14] H. Dette, M. Hallin, T. Kley, S. Volgushev, Of copulas, quantiles, ranks and spectra: an L_1-approach to spectral analysis, 2014, arXiv preprint arXiv:1111.7205v2.

[15] R. Chicheportiche, Non Linear Dependences in Finance (Ph.D. thesis), École Centrale Paris, 2013, Jun.
[16] E.G. Altmann, H. Kantz, Recurrence time analysis, long-term correlations, and extreme events, Phys. Rev. E 71 (5) (2005) 056106.
[17] M. Politi, T. Kaizoji, E. Scalas, Full characterization of the fractional Poisson process, Europhys. Lett. 96 (2) (2011) 20004.
[18] N. Sazuka, J.-I. Inoue, E. Scalas, The distribution of first-passage times and durations in FOREX and future markets, Physica A 388 (14) (2009) 2839–2853.
[19] M. Kac, On the notion of recurrence in discrete stochastic processes, Bull. Amer. Math. Soc. 53 (10) (1947) 1002–1010.
[20] A. Corral, Renormalization-group transformations and correlations of seismicity, Phys. Rev. Lett. 95 (2005) 8501.
[21] A. Corral, Point-occurrence self-similarity in crackling-noise systems and in other complex systems, J. Stat. Mech. Theory Exp. 2009 (01) (2009) P01022.
[22] G. Molchan, Interevent time distribution in seismicity: a theoretical approach, Pure Appl. Geophys. 162 (6) (2005) 1135–1150.
[23] G.F. Newell, M. Rosenblatt, Zero crossing probabilities for Gaussian stationary processes, Ann. Math. Stat. 33 (4) (1962) 1306–1313.
[24] P. Olla, Return times for stochastic processes with power-law scaling, Phys. Rev. E 76 (1) (2007) 011122.
[25] J. Ludescher, C. Tsallis, A. Bunde, Universal behaviour of interoccurrence times between losses in financial markets: An analytical description, Europhys. Lett. 95 (6) (2011) 68002.
[26] J. Perelló, J. Masoliver, J.-P. Bouchaud, Multiple time scales in volatility and leverage correlations: a stochastic volatility model, Appl. Math. Finance 11 (1) (2004) 27–50.
[27] K. Yamasaki, L. Muchnik, S. Havlin, A. Bunde, H.E. Stanley, Scaling and memory in volatility return intervals in financial markets, Proc. Natl. Acad. Sci. USA 102 (26) (2005) 9424–9428.
[28] F. Wang, K. Yamasaki, S. Havlin, H.E. Stanley, Scaling and memory of intraday volatility return intervals in stock markets, Phys. Rev. E 73 (2006) 6117.
[29] R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, C.E. Elger, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state, Phys. Rev. E 64 (2001) 061907.
[30] M. Boguná, J. Masoliver, Conditional dynamics driving financial markets, Eur. Phys. J. B 40 (3) (2004) 347–352.
[31] F. Ren, W.-X. Zhou, Recurrence interval analysis of high-frequency financial returns and its application to risk estimation, New J. Phys. 12 (7) (2010) 5030.
[32] W.-J. Xie, Z.-Q. Jiang, W.-X. Zhou, Extreme value statistics and recurrence intervals of NYMEX energy futures volatility, 2012. Arxiv preprint arXiv:q-fin.ST/1211.5502, November.