Jo, Hang Hyun

**Analytically solvable autocorrelation function for weakly correlated interevent times**

*Published in:*
Physical Review E

*DOI:*
10.1103/PhysRevE.100.012306

Published: 15/07/2019

*Document Version*
Publisher's PDF, also known as Version of record

*Please cite the original version:*
Analytically solvable autocorrelation function for weakly correlated interevent times

Hang-Hyun Jo

Asia Pacific Center for Theoretical Physics, Pohang 37673, Republic of Korea; Department of Physics, Pohang University of Science and Technology, Pohang 37673, Republic of Korea; and Department of Computer Science, Aalto University, Espoo FI-00076, Finland

(Received 3 January 2019; published 15 July 2019)

Long-term temporal correlations observed in event sequences of natural and social phenomena have been characterized by algebraically decaying autocorrelation functions. Such temporal correlations can be understood not only by heterogeneous interevent times (IETs) but also by correlations between IETs. In contrast to the role of heterogeneous IETs on the autocorrelation function, little is known about the effects due to the correlations between IETs. To rigorously study these effects, we derive an analytical form of the autocorrelation function for the arbitrary IET distribution in the case with weakly correlated IETs, where the Farlie-Gumbel-Morgenstern copula is adopted for modeling the joint probability distribution function of two consecutive IETs. Our analytical results are confirmed by numerical simulations for exponential and power-law IET distributions. For the power-law case, we find a tendency of the steeper decay of the autocorrelation function for the stronger correlation between IETs. Our analytical approach enables us to better understand long-term temporal correlations induced by the correlations between IETs.

DOI: 10.1103/PhysRevE.100.012306

I. INTRODUCTION

A variety of dynamical processes in natural and social phenomena have been described by a series of events or event sequences showing a non-Poissonian or bursty nature. Examples include solar flares [1], earthquakes [2,3], neuronal event sequences showing a non-Poissonian or bursty nature, and other works, modeling and numerical approaches were taken to investigate how strong correlations between IETs should be...
present to violate the scaling relations in Eq. (4) [17,29,30]. This situation clearly calls for a rigorous, analytical approach to the role of correlations between IETs in temporal correlations. For this, the correlations between IETs can be quantified by a memory coefficient $M$ [31] among others, such as local variation [32] or bursty trains [9]. The memory coefficient is defined as the Pearson correlation coefficient between two consecutive IETs, whose value for a sequence of $n$ IETs, i.e., $\{\tau_1, \ldots, \tau_n\}$, can be estimated by

$$M = \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{(\tau_i - \mu_1)(\tau_{i+1} - \mu_2)}{\sigma_1 \sigma_2},$$

where $\mu_1$ ($\mu_2$) and $\sigma_1$ ($\sigma_2$) are the average and the standard deviation of the first (last) $n-1$ IETs, respectively. Positive $M$ implies that the large (small) IETs tend to be followed by large (small) IETs. Negative $M$ indicates the opposite tendency, while $M = 0$ means the uncorrelated IETs. We mainly focus on the case with $M > 0$, based on the empirical observations [31,33–35].

To rigorously study the effects of correlations between IETs on the autocorrelation function, we derive an analytical form of the autocorrelation function for arbitrary $P(\tau)$ and for small $M$, i.e., in the case with weakly correlated IETs, where the Farlie-Gumbel-Morgenstern copula [36,37] is adopted for modeling the joint probability distribution function of two consecutive IETs. Our analytical results are numerically confirmed for both exponential and power-law IET distributions. In particular, for the power-law case, we find the steeper decay of the first (last) IETs for $M > 0$, based on the empirical observations [31,33–35].

We mainly focus on the case with $M > 0$, based on the empirical observations [31,33–35].

II. RESULTS

A. Analysis

We analyze the autocorrelation function $A(t_d)$ in Eq. (1). Since $A(t_d = 0) = 1$ is obvious, we consider the case with $t_d > 0$ unless otherwise stated. Note that for a Poisson process, i.e., without any memory effects in it, $A(t_d = 0) = 0$ for all $t_d > 0$. The event sequence $x(t)$ has the value of 1 at the moment the event occurred, 0 otherwise. Each event is assumed to have a duration of 1. Since events may overlap with each other due to their duration, we set the lower bound of IETs as 1, i.e., $\tau_{\min} = 1$, for the sake of simplicity. Then for an event sequence with $n$ events during the time period $T$, we get $\lambda \equiv \langle x(t) \rangle = n/T = 1/\mu$, with $\mu$ denoting the mean IET.

Using this $\lambda$, one can write

$$\langle x(t) x(t + t_d) \rangle = \lambda \sum_{k=1}^{\infty} P_k(t_d),$$

where $P_k(t_d)$ is the probability that two events that occurred at times $t$ and $t + t_d$ are separated by exactly $k$ interevent times (IETs) for $k = 1, 2, \ldots$. Using the joint probability distribution function (PDF) of $k$ consecutive IETs, denoted by $P(\tau_1, \ldots, \tau_k)$, one gets

$$P_k(t_d) = \prod_{i=1}^{k} \int_{0}^{\infty} d\tau_i P(\tau_1, \ldots, \tau_k) \delta \left( t_d - \sum_{i=1}^{k} \tau_i \right).$$

where $\delta(\cdot)$ is a Dirac delta function. Then the autocorrelation function in Eq. (1) can be rewritten as

$$A(t_d) = \frac{\sum_{k=1}^{\infty} P_k(t_d) - \lambda}{1 - \lambda},$$

where we have used $\langle x(t)^2 \rangle = \langle x(t) \rangle = \lambda$ as $x(t) = 0$, 1.

Since we only consider the correlations between two consecutive IETs, $P(\tau_1, \ldots, \tau_k)$ in Eq. (7) can be factorized in terms of joint PDFs of two consecutive IETs, i.e., $P(\tau_i, \tau_{i+1})$ for $i = 1, \ldots, n - 1$. More precisely, by assuming that an IET, $\tau_{i+1}$, is conditioned only by its previous IET, $\tau_i$, namely

$$P(\tau_{i+1} | \tau_i, \tau_{i-1}, \ldots) = P(\tau_{i+1} | \tau_i),$$

one obtains

$$P(\tau_1, \ldots, \tau_k) = \prod_{i=1}^{k-1} P(\tau_i, \tau_{i+1}) \prod_{i=2}^{k} P(\tau_i).$$

For modeling $P(\tau_i, \tau_{i+1})$, we adopt a Farlie-Gumbel-Morgenstern (FGM) copula among others [36,37], because the FGM copula is simple and analytically tractable, despite the range of correlation being somewhat limited, which will be discussed later. The FGM copula is originally defined as a function $C$ joining a bivariate cumulative distribution function (CDF) to the one-dimensional marginal CDFs such that

$$G(x_1, x_2) = C(u_1(x_1), u_2(x_2)) = u_1 u_2 (1 + r(1 - u_1)(1 - u_2)),\quad (11)$$

where $u_1$ ($u_2$) is a CDF of variable $x_1$ ($x_2$), and $r$ controls the correlation between $x_1$ and $x_2$ [36,37]. The bivariate PDF of $x_1$ and $x_2$ is obtained by

$$\frac{\partial^2 G(x_1, x_2)}{\partial x_1 \partial x_2} = P_1(x_1)P_2(x_2)(1 + r(2u_1 - 1)(2u_2 - 1)),\quad (12)$$

where $P_1(x_1)$ and $P_2(x_2)$ denote PDFs of $x_1$ and $x_2$, respectively. This FGM copula has been applied, e.g., for modeling the bivariate luminosity function of galaxies [37] and for health care data analysis [38].

The joint PDF of two consecutive IETs based on the FGM copula is written as

$$P(\tau_i, \tau_{i+1}) = P(\tau_i)P(\tau_{i+1})[1 + r f(\tau_i)f(\tau_{i+1})],\quad (13)$$

where

$$f(\tau) \equiv 2F(\tau) - 1,\quad F(\tau) \equiv \int_0^{\tau} d\tau' P(\tau').\quad (14)$$

Here $P(\tau_i)$ and $P(\tau_{i+1})$ are assumed to have the same functional form. The range of the parameter $r$ is given as $|r| \leq 1$ because $|f(\tau)| \leq 1$ from $0 \leq F(\tau) \leq 1$ and $P(\tau_i, \tau_{i+1}) \geq 0$. To relate $r$ to $M$ in Eq. (5), we redefine $M$ as

$$M = \frac{\langle \tau_{i+1} \rangle - \mu^2}{\sigma^2},$$

where $\langle \tau_{i+1} \rangle = \sum_{k=1}^{\infty} kP_k(t_d)$.
where
\[
\langle \tau_i \tau_{i+1} \rangle \equiv \int_0^\infty dt_i \int_0^\infty dt_{i+1} \tau_i \tau_{i+1} P(\tau_i, \tau_{i+1}),
\]
and \( \mu \) and \( \sigma \) are the mean and standard deviation of IETs, respectively. Using Eq. (13) we get
\[
M = \frac{r}{\sigma^2} \left[ \int_0^\infty d\tau \tau P(\tau)f(\tau) \right]^2 \equiv ar.
\]
The ratio \( a \) between \( M \) and \( r \) is determined only by \( P(\tau) \), irrespective of the correlations between IETs. Note that the upper bound of \( a \) is 1/3 for any \( P(\tau) \), hence \( |M| \leq 1/3 \) [39]. Due to this bound, applications of the FGM copula are limited to weakly correlated cases.

By plugging Eq. (13) into Eq. (10), we get
\[
P(\tau_1, \ldots, \tau_k) = \prod_{i=1}^k P(\tau_i) \left[ 1 + r \sum_{i=1}^{k-1} f(\tau_i) f(\tau_{i+1}) + O(r^2) \right],
\]
which enables us to calculate the Laplace transform of \( P_2(t_d) \) in Eq. (7):
\[
\tilde{P}_2(s) \equiv \int_0^\infty dt_1 P_2(t_d)e^{-st_1} = \tilde{P}(s)^k + r(k-1)\tilde{P}(s)^{k-2}\tilde{Q}(s)^2 + O(r^2),
\]
with
\[
\tilde{P}(s) \equiv \int_0^\infty d\tau P(\tau)e^{-s\tau},
\]
\[
\tilde{Q}(s) \equiv \int_0^\infty d\tau P(\tau)f(\tau)e^{-s\tau}.
\]
Then we obtain up to the first order of \( r \)
\[
\sum_{i=1}^\infty \tilde{P}_2(s) \approx \frac{\tilde{P}(s)}{1 - \tilde{P}(s)} + \frac{r^2\tilde{Q}(s)^2}{[1 - \tilde{P}(s)]^2} + O(r^2). \tag{23}
\]
The calculation of the higher-order terms of \( r \) is straightforward. By taking the inverse Laplace transform of Eq. (23) and plugging it into Eq. (8), we finally get the autocorrelation function as a function of \( M \) for the arbitrary form of \( P(\tau) \), which is denoted by \( A_M(t_d) \) hereafter.

\[\text{FIG. 1. Case with the exponential IET distribution in Eq. (24): Simulation results of the autocorrelation function } A_M(t_d) \text{ for various values of } \mu \text{ and } M \text{ (symbols)} \text{ are collapsed when rescaled properly. They are in good agreement with the analytical result up to the first order of } M \text{ in Eq. (26) (black dotted curve). Each point and its standard error were obtained over } 10^4 \text{ event sequences of } n = 5 \times 10^4.\]

B. Exponential IET distribution

One can consider the case with exponentially distributed IETs that are correlated with each other. Despite the fact that it is hard to find real-world examples of this case, we study this case because it is a good testbed for our analytical framework. More precisely, we use the following form of \( P(\tau) \) with the mean \( \mu \gg \tau_{\text{min}} = 1:\)
\[
P(\tau) = \mu^{-1} e^{-\tau/\mu}, \tag{24}
\]
by which one gets \( a = 1/4 \) in Eq. (17), hence \( r = 4M \). From Eq. (24), one gets
\[
\tilde{P}(s) = \frac{1}{\mu s + 1}, \quad \tilde{Q}(s) = \frac{-\mu s}{(\mu s + 1)(\mu s + 2)}. \tag{25}
\]
Plugging Eq. (25) into Eq. (23) and using Eq. (8), we analytically derive the autocorrelation function up to the first order of \( M \) as
\[
A_M(t_d) \approx \frac{4M_1 e^{-2\lambda t_d/\mu}}{\mu(\mu - 1)} + O(M^2), \tag{26}
\]
where \( \lambda = 1/\mu \) has been used. Note that the first term on the right-hand side in Eq. (26) can be written as \([4M/(\mu - 1)]g(t_d/\mu)\) with \( g(x) = xe^{-2x} \), implying that \( A_M(t_d) \) for various values of \( \mu \) and \( M \) can be collapsed when rescaled properly.

For the numerical validation of our analytical result, we introduce an algorithm for generating the event sequence using the FGM copula provided that \( P(\tau) \) and \( M \) are given, which is called the copula-based algorithm [40]: To generate a sequence of \( n \) IETs, i.e., \( \{\tau_1, \ldots, \tau_n\} \), the first IET \( \tau_1 \) is drawn from \( P(\tau) \) and the second IET \( \tau_2 \) is drawn from the conditional PDF \( P(\tau_2|\tau_1) = \tau(\tau_2)/P(\tau_1) \), where \( P(\tau_1, \tau_2) \) is modeled by the FGM copula in Eq. (13). Then \( \tau_i \) for \( i = 3, \ldots, n \) are sequentially drawn. Once the sequence of \( n \) IETs is ready, the timings of \( n + 1 \) events are set to be \( t_0 = 0 \) and \( t_i = \sum_{j=0}^{i-1} \tau_j \) for \( i = 1, \ldots, n \); the event sequence \( x(t) \) has the value of 1 for \( t \in [t_0, \ldots, t_n] \), otherwise \( x(t) = 0 \). This \( x(t) \) is then used to calculate the autocorrelation function in Eq. (1).

As shown in Fig. 1, the simulation results using Eq. (24) for various values of \( \mu \) and \( M \) are in good agreement with our analytical result in Eq. (26).
C. Power-law IET distribution

To be more realistic, we consider a power-law IET distribution with an exponential cutoff:
\[
P(\tau) = \frac{\tau^{\alpha-1}}{\Gamma(1 - \alpha, 1/\tau_c)} \tau^{-\alpha} e^{-\tau/\tau_c} \theta(\tau - 1),
\]
where \(\alpha\) and \(\tau_c\) denote the power-law exponent and exponential cutoff, respectively. \(\Gamma(\cdot, \cdot)\) is an upper incomplete Gamma function and \(\theta(\cdot)\) is a Heaviside step function, implying \(\tau_{\text{min}} = 1\). We also set \(\tau_c = 10^6\) for the rest of the paper, which is sufficiently large for studying the scaling behavior of the autocorrelation function. With this setup we numerically obtain the value of \(\alpha\) in Eq. (17), e.g., \(\alpha \approx 0.0039\) for \(\alpha = 1.4\) and \(\alpha \approx 0.0033\) for \(\alpha = 2.7\), respectively.

Since the analysis of the autocorrelation function with Eq. (27) is not straightforward, we instead use a simple power-law function for the IET distribution as
\[
P(\tau) = (\alpha - 1) \tau^{-\alpha} \theta(\tau - 1),
\]
which allows us to study the scaling behavior of the autocorrelation function to some extent. From Eq. (28) one gets
\[
\tilde{P}(s) = (\alpha - 1) s^{\alpha-1} \Gamma(1 - \alpha, s),
\]
\[
\tilde{Q}(s) = \tilde{P}(s) - 2(\alpha - 1) s^{2\alpha-2} \Gamma(2 - 2\alpha, s).
\]
We first analyze the case with \(1 < \alpha < 2\). In the asymptotic limit of \(s \to 0\), one obtains
\[
\tilde{P}(s) \approx 1 + b_1 s^{\alpha-1} + c_1 s + O(s^2),
\]
\[
\tilde{Q}(s) \approx b_2 s^{\alpha-2} - (c_1 - c_2) s + O(s^2),
\]
where \(\alpha \neq 3/2\)
\[
b_1 = \Gamma(1 - \alpha)(\alpha - 1), \quad c_1 = \frac{\alpha - 1}{2 - \alpha},
\]
\[
b_2 = \Gamma(2 - 2\alpha)(2\alpha - 2), \quad c_2 = \frac{2\alpha - 2}{3 - 2\alpha}.
\]
From Eqs. (8) and (23), and with \(\lambda = 0\) due to the diverging \(\mu\), we get for \(\alpha \neq 3/2\)
\[
A_M(t_d) \approx B_1 I_d^{(2-\alpha)} + B_2 t_d^{(3-\alpha)} + \ldots
\]
\[
\quad + \frac{M}{\alpha} (C_1 I_d^{(\alpha-1)} + C_2 t_d^{(2\alpha-3)} + \ldots) + O(M^2),
\]
where
\[
B_1 = \frac{-1}{b_1 \Gamma(1 - \alpha)}, \quad B_2 = \frac{c_1}{b_2^{2} \Gamma(2\alpha - 3)},
\]
\[
C_1 = \frac{-2b_2}{b_1 \Gamma(1 - \alpha)}, \quad C_2 = \frac{-2c_2}{b_1 \Gamma(2 - \alpha)}.
\]

In the case with uncorrelated IETs, i.e., \(M = 0\), the leading term of \(t_d^{(3-\alpha)}\) leads to the well-known scaling relation of \(\alpha + \gamma = 2\) for \(1 < \alpha < 2\) in Eq. (4).

The above analytical result in Eq. (33) is to be validated by the simulation results using Eq. (27). For the uncorrelated IETs, \(A_0(t_d)\) for \(\alpha = 1.4\) is calculated from the event sequences generated using the copula-based algorithm, as depicted in Fig. 2(a). The simulation result of \(A_0(t_d)\) turns out to be in good agreement with our analytical result in Eq. (33) with \(M = 0\) for several decades of \(t_d\). To confirm the effects due to the correlations between IETs, \(A_M(t_d)\) is numerically obtained for \(\alpha = 1.4\) and \(M = 0.002\) (i.e., \(r = M/\alpha \approx 0.52\)). Then we calculate its difference from the uncorrelated case, i.e., \(A_{0.002}(t_d) - A_0(t_d)\), which is found to be comparable to the analytical result up to the first order of \(M\) in Eq. (33); see Fig. 2(b).

Next, we analyze the case with \(2 < \alpha < 3\), where \(\mu\) is finite and \(\lambda = 1/\mu = -1/c_1\), to obtain for \(\alpha \neq 5/2\)
\[
A_M(t_d) \approx B'_1 I_d^{(\alpha-2)} + B'_2 t_d^{(\alpha-2)} + \ldots
\]
\[
\quad + \frac{M}{\alpha} (C'_1 I_d^{(\alpha-1)} + C'_2 t_d^{(\alpha-3)} + \ldots) + O(M^2),
\]
where
\[
B'_1 = \frac{b_1}{c_1(c_1 + 1) \Gamma(3 - \alpha)}, \quad B'_2 = \frac{-b_2^2}{c_1(c_1 + 1) \Gamma(5 - 2\alpha)},
\]
\[
C'_1 = \frac{2(c_1 - c_2) c_2 b_1}{c_1(c_1 + 1) \Gamma(2 - \alpha)}, \quad C'_2 = \frac{(3c_2 - 2c_1) c_2 b_1^2}{c_1(c_1 + 1) \Gamma(4 - 2\alpha)}.
\]
For the case with \(M = 0\), the leading term of \(I_d^{(\alpha-2)}\) leads to the well-known scaling relation of \(\alpha + \gamma = 2\) for \(2 < \alpha < 3\) in Eq. (4).

We find that the simulation results of \(A_0(t_d)\) and of the difference of \(A_{0.002}(t_d) - A_0(t_d)\) for \(\alpha = 2.7\) from the event sequences generated using the copula-based algorithm are comparable to our analytical result in Eq. (34), as evidenced in Figs. 2(c) and 2(d). Note that \(M = 0.002\) means \(r = M/\alpha \approx 0.60\). The discrepancy for the difference of \(A_{0.002}(t_d) - A_0(t_d)\) between the analytical and simulation results might be attributed to the finite \(\tau_c\) and/or \(n\).

Finally, we discuss the effect of \(\alpha\) on the overall decaying behavior of \(A_M(t_d)\). We make two observations in Eq. (33):
implying that the scaling relations in Eq. (4) can be easily violated by the correlations between IETs.

III. CONCLUSION

To investigate the effects of correlations between interevent times (IETs) on the autocorrelation function, we have derived the analytical form of the autocorrelation function for the arbitrary IET distribution $P(\tau)$ and for small values of the memory coefficient $M$, i.e., in the case with weakly correlated IETs, where the Farlie-Gumbel-Morgenstern copula [36,37] is adopted for modeling the joint probability distribution function of two consecutive IETs. For numerical validation, the event sequences are generated using the copula-based algorithm [40], by which IETs can be drawn sequentially only conditioned by their previous IETs. For both exponential and power-law IET distributions, we find that the simulation results of autocorrelation functions are in good agreement with the corresponding analytical solutions.

In particular, for the power-law case, we find that the stronger correlation between IETs with larger $M$ leads to the steeper decay of the autocorrelation function. In other words, the apparent decaying exponent $\gamma$ is found to increase with $M$. Our finding sheds light on the effects of correlations between IETs on other measures for temporal correlations, too, such as the Hurst exponent $H$ and the scaling exponent of the power spectral density $\eta$, considering their interdependence [7,8,15,16]. We also expect to better understand the differences between the empirical autocorrelation functions and those calculated for the randomized event sequences [9,16] based on our results. Finally, our results also support the previous numerical finding on the increasing tendency of $\gamma$ for the stronger correlation between IETs [17], where the correlations between IETs have been controlled by the power-law exponent of bursty train size distributions. Here we note that the bursty train size distribution and $M$ have been related to each other [41].

We remark that our analytical approach has the following limits: (i) We have considered only the correlations between two consecutive IETs based on the empirical findings, while the correlations between an arbitrary number of consecutive IETs have also been empirically observed in terms of heavy-tailed distributions of bursty train sizes [9,10,33]. This requires us to devise the more general analytical approach than ours as a future work. (ii) The FGM copula allows only relatively weak correlations between IETs, requiring us to consider other copulas for the cases with the stronger correlation between IETs [36]. Despite such limits, our analytical approach can help us to better understand the long-term temporal correlations ubiquitously observed in various natural and social phenomena, as little is known about the effects of the correlations between IETs on the long-term temporal correlations.

ACKNOWLEDGMENTS

The author thanks Takayuki Hiraoaka for fruitful discussions and acknowledges financial support from the Basic Science Research Program through the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (Grant No. NRF-2018R1D1A1A09081919).


