Temporal fluctuation scaling in nonstationary counting processes

Shinsuke Koyama
Department of Statistical Modeling, The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
(Dated: May 7, 2014)

The fluctuation scaling law has been observed in a wide variety of phenomena. For counting processes describing the number of events occurred during time intervals, it is expressed as a power function relationship between the variance and the mean of the event count per unit time, the characteristic exponent of which is obtained theoretically in the limit of long duration of counting windows. Here I show that the scaling law effectively appears even in a short timescale in which only a few events occur. Consequently, the counting statistics of nonstationary event sequences are shown to exhibit the scaling law as well as the dynamics at temporal resolution of this timescale. I also propose a method to extract in a systematic manner the characteristic scaling exponent from nonstationary data.

The fluctuation scaling law has been observed in many natural and man-made systems. It was originally found by Taylor in ecological systems as an empirical power function relationship between the variance and the mean of the number of individuals of a species per unit area [1]. The scaling relationship has been demonstrated in other fields such as transmission of infectious diseases, cancer metastasis, chromosomal structure and traffic in transportation networks [2–6], showing a universality of the law.

This letter focuses particularly on the fluctuation scaling law in counting processes. A counting process is a stochastic process \( \{ N_t; t \geq 1 \} \) describing the number of events occurred in the interval \((0, t] \), which is used for modeling a wide variety of phenomena such as occurrence of earth quake, photon counting and neural spike trains [7–9]. The fluctuation scaling law for the counting process considered here states that the variance of \( N_t \) per unit time is a power function of the mean of \( N_t \) per unit time,

\[
\text{Var}(N_t)/t \propto [E(N_t)/t]^\beta. \tag{1}
\]

Since a random process leads to a Poisson process, \( \beta = 1 \) becomes an indicator of randomness: every deviation from randomness indicates a deviation from this relationship.

To compute the mean and the variance of \( N_t \), it is usually taken a counting window of long duration \( t \gg 1 \) in which a large number of events occur. However, the scaling law [1] generally depends on the duration of the counting window, or on the average number of events in the window. In the limit of \( t \to 0 \), for instance, it can be shown that the scaling relation with arbitrary exponent vanishes but \( \beta \) approaches unity, which is essentially the same as the fact that the Fano factor \( F(t) := \text{Var}(N_t)/E(N_t) \) for any regular point process approaches unity for \( t \to 0 \) [10]. It is therefore of interest how many events on average in the counting window is enough to observe the scaling law, the exponent of which characterizes the ‘intrinsic’ variability of occurrence of events.

This question is important particularly for nonstationary sequences of events. In nervous systems, for example, the firing rate is typically modulated with timescale of tens to hundreds milliseconds, in which only a few events (spikes) occur [11]. Since neurons operate in such a short timescale, it is important to ask if the counting statistics exhibits the scaling law with only a few events.

Here, I show by assuming renewal processes that the scaling law in the counting statistics appears even in a short counting window, in which only a few events on average occur. I also propose a method to extract in a systematic manner the characteristic scaling exponent from nonstationary sequences of events. The ability of the proposed method is demonstrated with data simulated by a leaky integrate-and-fire neuron model.

I begin with the fluctuation scaling law for stationary renewal processes. Let \( X \geq 0 \) be an interevent interval, and \( m = E(X) \) and \( s^2 = \text{Var}(X) \) be its mean and variance, respectively. Suppose that the variance has a power function relation with \( m \) as

\[
s^2 = \kappa m^\alpha. \tag{2}
\]

The scaling exponent \( \alpha \) characterizes the ‘intrinsic’ dispersion of occurrence of events. For a Poisson (random) process, \( \alpha = 2 \). On the other hand, \( \alpha > 2 \) \((< 2)\) implies the tendency for the timing of event occurrence to be over (under) dispersed for large means, and under (over) dispersed for small means.

Let \( N_t \) be the number of events occurred in the counting window \((0, t] \). For \( t \gg 1 \), \( N_t \) asymptotically follows the Gaussian distribution with mean \( t/m \) and variance \( s^2t/m^3 \) [12]. Then, if the interval statistics has the scaling property [2], the variance of \( N_t \) per unit time is asymptotically scaled by the mean of \( N_t \) per unit time (i.e., the rate) as

\[
\frac{\text{Var}(N_t)}{t} = \kappa \left[ \frac{E(N_t)}{t} \right]^\beta, \tag{3}
\]
where $\beta = 3 - \alpha$ for $t \gg 1$. Note that the scaling law depends on the duration $t$ of the counting window. In theory, the exponent $\beta = 3 - \alpha$ is achieved in the limit of $t \to \infty$, in which a sufficiently large number of events occur. On the other hand, $\beta$ approaches 1 for the limit of $t \to 0$.

One can construct an interevent interval density of renewal processes that possesses the scaling law by introducing a parametric probability density $f(z; \kappa)$ with unit mean and the variance $\kappa$, and rescaling it as

$$p(x; \lambda, \kappa, \alpha) = \lambda f(\lambda x; \Lambda^{2-\alpha} \kappa),$$

where $\lambda = 1/m$ is the frequency of occurrence of events. Here, the choice of $f(z; \kappa)$ is arbitrary: different choices of $f(z; \kappa)$ generate different families of probability densities that have the scaling law.

A nonstationary renewal process that exhibits the scaling law in the counting statistics can be constructed by generalizing the construction of nonstationary Poisson processes, the idea of which is given as follows. Consider a stationary Poisson process with unit rate defined on dimensionless time $\tau$. Then, the probability of occurring an event in a short interval $[s, s + ds]$ is given by $ds$. Let $\Lambda(t)$ be a rate of event occurrence on the real time $t$ and $\Lambda(t) = \int_0^t \lambda(u) du$ be the cumulative function of $\Lambda(t)$. By transforming the time $\tau$ into time $t$ with $s = \Lambda(t)$, one obtain a nonstationary Poisson process with time-dependent rate $\Lambda(t)$, in which the probability of occurring an event in a short interval $[t, t + dt]$ is $\Lambda(t) dt$. In the same manner, any renewal process with unit rate can be transformed by $s = \Lambda(t)$ into a nonstationary renewal process with the trial-averaged rate $\Lambda(t)$.

However, this transformation does not allow the variance of the event count per unit time to have the power function of the rate with arbitrary scaling exponent.

Hence, I propose a generalization of the transformation so that the variance and the mean of count per unit time obey the scaling law. Consider a renewal process with the interevent interval density $f(z; \kappa)$. The conditional rate, or the hazard function, of this process is given by

$$g(s; s_s, \kappa) = \frac{f(s; \kappa)}{1 - \int_{s_s}^s f(u; \kappa) du}.$$ \hspace{1cm} (5)

where $s_s(s < s)$ is the last event time preceding $s$. Analogously to Eq. (4), by rescaling the variance parameter $\kappa \to \lambda(t)^{2-\alpha} \kappa$ as well as the time $s = \Lambda(t)$, the conditional rate of the nonstationary renewal process is obtained as

$$r(t; t_s, \{\Lambda(t)\}, \kappa, \alpha) = \frac{\lambda(t) f(\Lambda(t) - \Lambda(t_s); \lambda(t)^{2-\alpha} \kappa)}{1 - \int_{t_s}^t \lambda(v) f(\Lambda(v) - \Lambda(t_s); \lambda(v)^{2-\alpha} \kappa) dv}.$$ \hspace{1cm} (6)

For $dt \ll 1$, Eq. (6) gives the conditional probability of occurring an event in $[t, t + dt]$, given the last event at $t_s$,

$$P(N_{t+dt} - N_t = 1; t_s, \{\Lambda(t)\}, \kappa, \alpha) \approx r(t; t_s, \{\Lambda(t)\}, \kappa, \alpha) dt.$$ \hspace{1cm} (7)

which can be used for simulating sequences of events.

With the basis of the model, I propose a method for estimating the scaling exponent $\alpha$ (and the coefficient $\kappa$) from data consisting of nonstationary sequences of events. The likelihood function of $(\alpha, \kappa)$, given a sequence of event times $\{t_i\} := \{t_1, \ldots, t_n\}$, is expressed with the conditional rate function $f(\cdot)$ as

$$l(\kappa, \alpha; \{t_i\}, \{\Lambda(t)\}) = \prod_{i=2}^n r(t_i; t_{i-1}, \{\Lambda(t)\}, \kappa, \alpha) \times \exp \left( -\int_{t_1}^{t_n} r(u; t_{N_s}, \{\Lambda(t)\}, \kappa, \alpha) du \right),$$ \hspace{1cm} (8)

where the exponential factor represents the probability of no event in each interevent interval. However, this transformation does not allow the variance of the event count per unit time to have the power function of the rate with arbitrary scaling exponent.

For $M$ independent and identically distributed trials $\{t_{ij}\}^M_{j=1} := \{t_{i1}, \ldots, t_{ijn}\}^M_{j=1}, n_j$ denoting the number of events in the $j$th trial, the likelihood function is simply given by the product of the likelihood function of single trials.

Using this, the parameters can be estimated in the following two steps. 1) Compute an estimate $\hat{\lambda}(t)$ of the trial-averaged rate function from $\{t_{ij}\}^M_{j=1}$, which can be obtained by a kernel density estimator with a Gaussian kernel whose bandwidth is determined by minimizing the expected mean squared error. 2) Substitute $\hat{\lambda}(t)$ and $\{t_{ij}\}^M_{j=1}$ into the likelihood function, and maximize it with respect to $(\alpha, \kappa)$.

To study the behavior of the statistical model, I propose a method for estimating the scaling exponent $\alpha$ and the coefficient $\kappa$ from data consisting of nonstationary sequences of events. The likelihood function of $(\alpha, \kappa)$, given a sequence of event times $\{t_i\} := \{t_1, \ldots, t_n\}$, is expressed with the conditional rate function $f(\cdot)$ as

$$l(\kappa, \alpha; \{t_i\}, \{\Lambda(t)\}) = \prod_{i=2}^n \lambda(t_i) f(\Lambda(t_i) - \Lambda(t_{i-1}); \lambda(t_i)^{2-\alpha} \kappa).$$ \hspace{1cm} (9)

For $M$ independent and identically distributed trials $\{t_{ij}\}^M_{j=1} := \{t_{i1}, \ldots, t_{ijn}\}^M_{j=1}, n_j$ denoting the number of events in the $j$th trial, the likelihood function is simply given by the product of the likelihood function of single trials. Using this, the parameters can be estimated in the following two steps. 1) Compute an estimate $\hat{\lambda}(t)$ of the trial-averaged rate function from $\{t_{ij}\}^M_{j=1}$, which can be obtained by a kernel density estimator with a Gaussian kernel whose bandwidth is determined by minimizing the expected mean squared error. 2) Substitute $\hat{\lambda}(t)$ and $\{t_{ij}\}^M_{j=1}$ into the likelihood function, and maximize it with respect to $(\alpha, \kappa)$.

To study the behavior of the statistical model, I propose a method for estimating the scaling exponent $\alpha$ and the coefficient $\kappa$ from data consisting of nonstationary sequences of events. The likelihood function of $(\alpha, \kappa)$, given a sequence of event times $\{t_i\} := \{t_1, \ldots, t_n\}$, is expressed with the conditional rate function $f(\cdot)$ as

$$l(\kappa, \alpha; \{t_i\}, \{\Lambda(t)\}) = \prod_{i=2}^n \lambda(t_i) f(\Lambda(t_i) - \Lambda(t_{i-1}); \lambda(t_i)^{2-\alpha} \kappa).$$ \hspace{1cm} (9)

For $M$ independent and identically distributed trials $\{t_{ij}\}^M_{j=1} := \{t_{i1}, \ldots, t_{ijn}\}^M_{j=1}, n_j$ denoting the number of events in the $j$th trial, the likelihood function is simply given by the product of the likelihood function of single trials. Using this, the parameters can be estimated in the following two steps. 1) Compute an estimate $\hat{\lambda}(t)$ of the trial-averaged rate function from $\{t_{ij}\}^M_{j=1}$, which can be obtained by a kernel density estimator with a Gaussian kernel whose bandwidth is determined by minimizing the expected mean squared error. 2) Substitute $\hat{\lambda}(t)$ and $\{t_{ij}\}^M_{j=1}$ into the likelihood function, and maximize it with respect to $(\alpha, \kappa)$.

To study the behavior of the statistical model, I propose a method for estimating the scaling exponent $\alpha$ and the coefficient $\kappa$ from data consisting of nonstationary sequences of events. The likelihood function of $(\alpha, \kappa)$, given a sequence of event times $\{t_i\} := \{t_1, \ldots, t_n\}$, is expressed with the conditional rate function $f(\cdot)$ as

$$l(\kappa, \alpha; \{t_i\}, \{\Lambda(t)\}) = \prod_{i=2}^n \lambda(t_i) f(\Lambda(t_i) - \Lambda(t_{i-1}); \lambda(t_i)^{2-\alpha} \kappa).$$ \hspace{1cm} (9)
large number of events occur, so that the central limit theorem can be applied \(^2\) The simulation result, however, shows that the scaling law with \(\beta = 3 - \alpha\) appears even with a few events.

To examine the nonstationary case, a periodic function \(\lambda(t) = 0.04 + 0.02 \sin \frac{\pi t}{50}\) is used for the time-dependent rate, and \(M = 10^4\) sequences of events are simulated using Eq. (7) in the time interval \(t \in (0, 1000]\) for each \(\alpha = 1, 2\) and 3. The time axis is divided into equally spaced, contiguous time windows, each of duration \(\Delta\), and the number of events in the \(i\)th window of the \(j\)th trial is counted and denoted by \(N_i^j\). The mean and the variance of the event count per unit time in the \(i\)th window are respectively computed as

\[
\hat{\lambda}_i = \frac{\sum_{j=1}^{M} N_i^j}{\Delta}, \tag{11}
\]

and

\[
\hat{\lambda}_i = \frac{\sum_{j=1}^{M} (N_i^j - \hat{\lambda}_i \Delta)^2}{\Delta}. \tag{12}
\]

The slope \(\beta\) is then computed by performing the linear regression of \{\log \hat{\lambda}_i\} on \{\log \hat{\lambda}_i\}. Figure 2 depicts \(\beta\) as a function of \(\Delta\), showing that \(\beta\) approaches 3 as \(\Delta \rightarrow 0\). (b) The log-log plot of \(\hat{\lambda}_i\) against \(\hat{\lambda}_i\), calculated with \(\Delta = 40\). The solid lines represent the theoretical slope \(\beta = 3 - \alpha\). (c) The time course of \(\lambda_i\) (top) and \(\hat{\lambda}_i\) (bottom) calculated with \(\Delta = 40\). (d) The same as in (c) but with \(\Delta = 1\).

The property of spike trains generated from the LIF model. The dynamics of the LIF model are represented by the equation \(^\frac{21}{21}\),

\[
\frac{dV(t)}{dt} = -\frac{V(t)}{\tau} + I(t), \tag{13}
\]

where \(V(t)\) is the membrane potential, \(\tau = 5\) is the membrane decay time constant, and \(I(t)\) represents the input current. When the membrane potential reaches the threshold \(v_{th} = 1\), an event (spike) is generated and the membrane potential is reset to 0 immediately.

For a stationary input current \(I(t) = \mu + \sigma \xi(t)\), where \(\xi(t)\) is a Gaussian white noise satisfying \(\langle \xi(t) \rangle = 0\) and \(\langle \xi(t) \xi(t') \rangle = \delta(t-t')\), the following two cases are examined in the simulations: (i) \(\mu\) varies from 0.15 to 0.25 while \(\sigma = 0.1\) is fixed, and (ii) \(\sigma\) varies from 0.2 to 0.6 while \(\mu = 0.05\) is fixed. For each set of parameter values \((\mu, \sigma)\), a sequence of \(10^3\) events is generated, from which the mean and the variance of interevent interval are calculated. It is found that the variance and the mean obey the scaling law, whose exponent obtained by the linear regression on a log-log scale is \(\alpha = 3.03\) for (i) and \(\alpha = 1.89\) for (ii) (Figure 3).
gously to the stationary case: (iii) the mean current has a periodic profile \(\mu(t) = 0.2 + 0.05 \sin \frac{2\pi t}{500}\), while the amplitude of the current fluctuation is constant \(\sigma(t) = 0.1\), and (iv) the amplitude of the current fluctuation is periodically modulated \(\sigma(t) = 0.4 + 0.2 \sin \frac{2\pi t}{500}\), while the mean current is constant \(\mu(t) = 0.05\). For each case, \(M = 10^4\) spike trains are simulated in the time interval \(t \in (0, 1000]\), from which \(\{\lambda_i\}\) and \(\{\hat{v}_i\}\) are computed by Eqs. (11)–(12) with \(\Delta = 40\). Figures 3a–c depict the result, showing that \(\hat{v}_i\) and \(\hat{\lambda}_i\) approximately obey the scaling law. The slope obtained by the linear regression of \(\log \hat{v}_i\) on \(\log \hat{\lambda}_i\) is \(\beta = 0.03\) for (iii) and \(\beta = 1.10\) for (iv), from which we see the approximate relation \(\beta \approx 3 - \alpha\). It is, thus, empirically confirmed that the spike trains generated from the LIF model exhibit the scaling law, which qualitatively matches that of the statistical model (4) (Figures 2a, b, and c).

I finally examine if the proposed inference method can capture the scaling exponent \(\alpha\) directly from the nonstationary sequences of events generated by the LIF model. For each nonstationary input current (iii) and (iv), the inference method is applied to \(M\) spike trains simulated by the LIF model to obtain \(\hat{\alpha}\). Figure 3d plots \(\hat{\alpha}\) against the number of trials \(M\), which shows that the accuracy of the estimation is improved as \(M\) is increased. For example, the exponent is estimated from \(M = 20\) trials as \(\hat{\alpha} = 3.08 \pm 0.05\) for the case (iii) and \(\hat{\alpha} = 1.93 \pm 0.03\) for the case (iv), which are in good agreement with the values obtained in Figure 3a.

For summary, it was shown in this letter that assuming renewal processes, only a few events in counting windows are enough for the variance and the mean per unit time to exhibit the scaling law with the exponent \(\beta = 3 - \alpha\). As a result, the counting statistics of nonstationary event sequences display the scaling law as well as the dynamics at temporal resolution of this counting windows (Figures 2a and 3a). I also proposed a method based on the likelihood principle to extract the scaling exponent from nonstationary sequences of events, the ability of which was demonstrated with the data simulated by the LIF model.

The results of renewal processes can be generalized to nonrenewal processes directly. For nonrenewal processes, whose interval statistics has the scaling law (2), the asymptotic scaling relation in the counting statistics (3) remains unchanged except that the coefficient is modified (22). The generalization of the proposed inference method to the nonrenewal processes is also straightforward since the transformation of a stationary point process (5) to a nonstationary point process (6) is applicable to nonrenewal processes.

In nervous systems, neurons produce an action potential by integrating presynaptic inputs within tens milliseconds, in which typically only a few spikes come from each presynaptic neuron. This suggests that the scaling law in spike count effectively appears in the integration time, and thus may have an impact on information processing. Ma et al. (21) suggested a hypothesis that the Poisson-like statistics in the responses of populations of cortical neurons may represent probability distributions over the stimulus and implement Bayesian inferences. An important property in their hypothesis is that the variance of spike count is proportional to the mean spike count, which corresponds to \(\beta = 1\) (or \(\alpha = 2\) in the interval statistics) in our formulation. It is worth pointing out that \(\beta \approx 1\) is observed in the simulations of the LIF neuron with the fluctuating current input (Figure 3d), red), which can be realized by balanced excitatory and inhibitory synaptic inputs observed in the cortex (22, 29).

On the other hand, from in vivo recordings, Troy and Robson (27) found that steady discharges of retinal ganglion cells, in response to stationary visual patterns, exhibit the scaling law in the interval statistics (2) with exponent \(\alpha \approx 3\). This exponent is also observed in the simulations using multi-compartment models of retinal ganglion cells (28) as well as the LIF model with the current input whose mean is modulated (Figure 3c, blue).

It is therefore speculated that the scaling exponent \(\alpha\) may reflect the intrinsic mechanisms of neuronal discharge or the internal dynamics of networks (30). and may

FIG. 3: The result of the LIF model. In all the figures, blue indicates that the mean current \(\mu\) varies, while red indicates that the amplitude of current fluctuation \(\sigma\) varies. (a) The log-log plot of the variance against the mean of interevent intervals when the stationary input currents are injected. The solid lines represent the slopes obtained by the linear regression on the log-log scale. (b) The log-log plot of \(\hat{v}_i\) against \(\hat{\lambda}_i\) calculated with \(\Delta = 40\) when the nonstationary current inputs are injected. The solid lines represent the slopes obtained by the linear regression on the log-log scale. (c) The time courses of \(\hat{\lambda}_i\) (top) and \(\hat{v}_i\) (bottom) calculated with \(\Delta = 40\). (d) The estimate \(\hat{\alpha}\) as a function of the number of trials \(M\). The standard deviations are computed with 100 realizations.
be related to schemes for neural computation the nervous systems employ. The proposed method offers a systematic way to extract the characteristic scaling exponent from experimental data.

This research was supported by JSPS KAKENHI Grant Number 24700287.

[1] L. R. Taylor, Nature 189, 732 (1961).
[2] R. M. Anderson and R. M. May, Nature 333, 514 (1989).
[3] W. S. Kendal and P. Frost, J. Natl. Cancer Inst. 79, 1113 (1987).
[4] W. S. Kendal, BMC Evol. Biol. 4, 3 (2004).
[5] A. Fronczak and P. Fronczak, Phys. Rev. E 81, 066112 (2010).
[6] M. A. de Menezes and A.-L. Barabasi, Phys. Rev. Lett. 92, 028701 (2004).
[7] Y. Ogata, J. Amer. Statist. Assoc. 83, 9 (1988).
[8] M. O. Scully and M. S. Zubairy, Quantum Optics (Cambridge University Press, 1997).
[9] D. H. Johnson, J. Comp. Neurosci. 3, 275 (1996).
[10] M. C. Teich et al., J. Opt. Soc. Am. A 14, 529 (1997).
[11] B. J. Richmond et al., J. Neurophysiol. 57, 132 (1987).
[12] D. R. Cox, Renewal Theory (Chapman and Hall, London, 1962).
[13] M. Berman, Biometrika 68, 143 (1981).
[14] R. Barbieri et al., J. Neurosci. methods 105, 25 (2001).
[15] S. Koyama and R. E. Kass, Neural Comp. 20, 1776 (2008).
[16] J. W. Pillow, in Neural Information Processing Systems, edited by Y. Bengio et al., Vol. 22, 1473, (2008).
[17] D. J. Daley and D. Vere-Jones, An Introduction to the Theory of Point Processes, Vol. 1 (Springer, New York, 2002).
[18] R. E. Kass and V. Ventura, Neural Comp. 13, 1713 (2001).
[19] H. Shimazaki and S. Shinomoto, J. Comp. Neurosci. 29, 171 (2010).
[20] C. Koch, Biophysics of Computation: Information Processing in Single Neurons (Oxford University Press, Oxford, 1999).
[21] W. J. Ma et al., Nature Neurosci. 9, 1432 (2006).
[22] C. V. Vreeswijk and H. Sompolinsky, Science 274, 1724, (1996).
[23] D. J. Amit and N. Brunel, Cereb. Cortex 7, 237, (1997).
[24] A. Destexhe et al., Neurosci. 107, 13 (2001).
[25] Y. Shu, A. Hasenstaub and D. A. McCormick, Nature 423, 288 (2003).
[26] B. Haider et al., J. Neurosci. 26, 4535 (2006).
[27] J. Troy and J. Robson, Visual Neurosci. 9, 535 (1992).
[28] M. C. W. van Rossum, B. J. O’Brien and R. G. Smith, J Neurophysiol. 89, 2406 (2003).
[29] M. P. Nawrot, in Analysis of Parallel Spike Trains, edited by S. Grun and S. Rotter, Chap. 3 (Springer, New York, 2010).
[30] M. A. de Menezes and A.-L. Barabasi, Phys. Rev. Lett. 93, 068701 (2004).
[31] In fact, this transformation results in the variance of the event count being proportional to the mean, Var(N_t) = \kappa E(N_t) for t \gg 1.
[32] For stationary point processes with serially correlated interevent intervals and the scaling relation (2), it holds for t \gg 1 that Var(N_t)/t = \kappa(1 + 2\eta)E(N_t)/t^{3-\alpha} with \eta = \sum_{i=1}^{\infty} \eta_i, where \eta_i denotes the ith-order linear correlation coefficient for pairs of intervals that are separated by i − 1 intermediate intervals.\]