ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

How regular is neuronal activity?

Lubomir Kostal, Petr Lansky and Ondrej Pokora *

Institute of Physiology AS CR, v.v.i., Dept. of Computational Neuroscience, Videnska 1083, Praha 4, Czech Republic

Abstract. We propose and investigate two information-based measures of statistical dispersion of neuronal firing: the entropy-based dispersion and Fisher information-based dispersion. The measures are compared with the standard deviation. Although the standard deviation is used routinely, we show, that it is not well suited to quantify some aspects of dispersion that are often expected intuitively, such as the degree of randomness. The proposed dispersion measures are not entirely independent, although each describes the firing regularity from a different point of view.

1 Introduction

Information-based measures of signal regularity or randomness have recently gained popularity in various branches of science, see e.g., [1, 2]. In this paper, we construct dispersion-like quantities based on these information measures and apply them. In particular, we continue the effort initiated in [3, 4] by taking into account a variant of Fisher information, which has been employed also in different contexts [2, 5].

Although standard deviation is used ubiquitously and is almost synonymous to the "measure of statistical dispersion", we show, that it is not well suited to quantify some aspects of spiking regularity that are often expected intuitively [3, 6]. For example, the diversity or randomness of the interspike intervals is better described by entropy-based or Fisher information-based dispersions.

We show, that the main difference between the descriptions by means of entropy or Fisher information, lies in the fact that the Fisher information describes how "smooth" the distribution is, while the entropy describes how "even" it is. The "smoothness" and "evenness" might be at first considered as interchangeable, but we show in detail that it is not the case. The illustration of the proposed methods is provided on simple and frequently employed models of stationary neuronal activity, given by exponential, gamma and inverse Gaussian distributions of interspike intervals (ISI).

2 Methods

The probabilistic description of ISIs results from the fact, that the positions of spikes usually cannot be predicted exactly, only the probability that a spike occurs is given [7]. Therefore there is a common approach to the spiking neuronal activity by describing it as a stochastic point process. The prominent role among

*This work was supported by AV0Z50110509, Centre for Neuroscience LC554 and the Grant Agency of the Czech Republic project P103/11/0282.

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

these processes is played by so called stationary renewal point process. Then the ISIs can be exclusively described by the continuous positive random variable, denoted as T and this situation is considered here.

Statistical dispersion is a measure of "variability" or "spread" of the distribution of the random variable T. The measure has the same physical units as T. By far, the most common measure of dispersion is the standard deviation, σ, defined as the second central moment of the distribution. The corresponding *relative* dispersion is known as the coefficient of variation, c_v,

$$c_v = \frac{\sigma}{\mathrm{E}\,(T)}, \tag{1}$$

where $\mathrm{E}\,(T)$ is the mean value of T. Standard deviation (or c_v) measures essentially how off-centered is the distribution of T and its value is sensitive to outlying values [8, 9]. On the other hand, σ does not quantify how random, or unpredictable, are the outcomes of r.v. T. Namely, high value of σ (high variability) does not indicate that the possible values of T are distributed evenly [3].

The randomness of a probability distribution can be defined as the measure of "choice" of different outcomes that are possible. For discrete probability distributions such measure is provided by the Shannon entropy, which is known to be a unique measure consistent with some natural requirements [10]. For continuous variables, however, the value of Shannon entropy diverges, therefore the following measure of randomness was proposed in [11]

$$\sigma_h = \exp\left[-\int_0^\infty f(t)\,\ln f(t)\,dt\right], \tag{2}$$

where $f(t)$ is the probability density function of random variable T. The interpretation of σ_h relies on the asymptotic equipartition property theorem and the entropy power concept [12, 11]. Informally, the value of σ_h is bigger for those random variables, which generate more diverse (or unpredictable) realizations. Analogously to Eq. (1), we define the relative entropy-based dispersion coefficient, c_h, as

$$c_h = \frac{\sigma_h}{\mathrm{E}\,(T)}. \tag{3}$$

The Fisher information measures the minimum possible error in estimating a parameter of a distribution. In a special case of the location parameter, the Fisher information $J(T)$ does not depend on the parameter itself, and can be expressed directly as a functional of the density $f(t)$ [12, p.671],

$$J(T) = \int_0^\infty \left[\frac{\partial \ln f(t)}{\partial t}\right]^2 f(t)\,dt. \tag{4}$$

The value of $J(T)$ is small for smoothly-shaped probability densities. Any locally steep slope or the presence of modes in the shape of $f(t)$ increases $J(T)$ [2]. Due

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

to the derivative in Eq. (4), certain regularity conditions are required on $f(t)$. In this paper, we assume that $f(t)$ is continuously differentiable for all $t > 0$ and furthermore $f(0) = f'(0) = 0$ [13]. The units of $J(T)$ correspond to the inverse of the squared units of T, therefore we propose the Fisher information-based dispersion measure, σ_J, as

$$\sigma_J = \frac{1}{\sqrt{J(T)}}. \tag{5}$$

Any "non-smoothness" in the shape of $f(t)$ decreases σ_J. In analogy with Eqns. (1) and (3) the relative dispersion coefficient c_J can also be defined,

$$c_J = \frac{\sigma_J}{\mathrm{E}\,(T)}. \tag{6}$$

3 Results

We choose three widely employed statistical models of ISIs: gamma [14, 15], inverse Gaussian [15, 16] and lognormal distributions [15], and analyze them by means of the three described dispersion coefficients c_v, c_h and c_J.

The probability density function of the gamma distribution, parametrized by shape parameter k and scale parameter θ, is

$$f(t) = \frac{t^{k-1}\ \exp\{-t/\theta\}}{\Gamma(k)\ \theta^k}, \tag{7}$$

where $\Gamma(z)$ is the gamma function. The mean value of the distribution is $\mathrm{E}\,(T) = k\,\theta$ and the coefficient of variation is equal to

$$c_v = 1/\sqrt{k}. \tag{8}$$

By parametrizing the density (7) by c_v and substituting it into Eqns. (3) and (6) we obtain the entropy-based and Fisher information-based dispersion coefficients as functions of c_v,

$$c_h = c_v^2\,\Gamma(c_v^{-2})\exp\left\{\frac{1+(c_v^2-1)\Psi(c_v^{-2})}{c_v^2}\right\}, \tag{9}$$

$$c_J = c_v\sqrt{1-2c_v^2} \qquad \text{for} \quad 0 < c_v < \frac{1}{\sqrt{2}}, \tag{10}$$

where $\Psi(z) = \frac{d}{dz}\ln\Gamma(z)$ is the digamma function [17]. Note, that the gamma density is not differentiable at $t = 0$ for $c_v \geq 1/\sqrt{2}$, thus c_J is evaluated only for $0 < c_v < 1/\sqrt{2}$.

The inverse Gaussian distribution describes the spiking activity of a stochastic variant of the perfect integrate-and-fire neuronal model [7, 18]. Its probability density function, parametrized by mean, $\mu = \mathrm{E}\,(T)$, and scale parameter, σ, can be expressed as

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2 t^3}}\exp\left\{-\frac{(t-\mu)^2}{2\sigma^2\mu^2 t}\right\}. \tag{11}$$

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

The coefficient of variation is equal to

$$c_v = \sqrt{\mu\,\sigma^2} \tag{12}$$

and the other dispersion coefficients can be expressed as

$$c_h = \sqrt{2\pi}\,c_v \exp\left\{\frac{1}{2} - \frac{3\exp(c_v^{-2})\,K^{(1,0)}\left(\frac{1}{2}, c_v^{-2}\right)}{\sqrt{2\pi}\,c_v}\right\}, \tag{13}$$

$$c_J = \frac{\sqrt{2}\,c_v}{\sqrt{2 + 9c_v^2 + 21c_v^4 + 21c_v^6}}, \tag{14}$$

where $K^{(1,0)}(\nu, z)$ is the derivative of the modified Bessel function of the second kind, $K^{(1,0)}(\nu, z) = \frac{\partial}{\partial \nu}K(\nu, z)$ [17].

The lognormal probability density function, parametrized by mean, μ, and standard deviation, σ, of variable $\ln T$, is

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}\;t}\exp\left\{-\frac{(\ln t - \mu)^2}{2\sigma^2}\right\}. \tag{15}$$

Using this parametrization, the mean of the lognormal distribution is $\mathrm{E}\,(T) = \exp\left\{\mu + \sigma^2/2\right\}$ and the coefficient of variation is equal to

$$c_v = \sqrt{\exp\left(\sigma^2\right) - 1}. \tag{16}$$

The two other dispersion coefficients, expressed as functions of c_v, are

$$c_h = \sqrt{2\pi e}\,\sqrt{\frac{\ln(1 + c_v^2)}{1 + c_v^2}}, \tag{17}$$

$$c_J = \sqrt{\frac{\ln(1 + c_v^2)}{[1 + c_v^2]^3[1 + \ln(1 + c_v^2)]}}. \tag{18}$$

By using these three statistical descriptors of interspike intervals, we can show how similar or different are the measures given by Eqns. (1), (3) and (6). The dependence of c_h on c_v is shown in Fig. 1a, the dependence of c_J on c_v is shown in Fig. 1b. The dependencies are not linear (even not monotonous) and thus neither c_h nor c_J are equivalent to c_v. We see, that both c_h and c_J as functions of c_v show a "$\cap$" shape with maxima around $c_v \doteq 1$ (for c_h) and around $c_v \doteq 0.5$ (for c_J). Note, that the plots of c_J against c_v appear like a scaled version of the plots of c_h against c_v, with the relative positions of the curves preserved (to a certain extent). In particular, while c_h of the lognormal is always greater than c_h of the inverse Gaussian, the ordering is reversed for c_J when $c_v > 2.2$. Furthermore, the maxima of c_h and c_J occur for different c_v values, confirming that each of the proposed dispersion coefficients provides a different point of view.

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

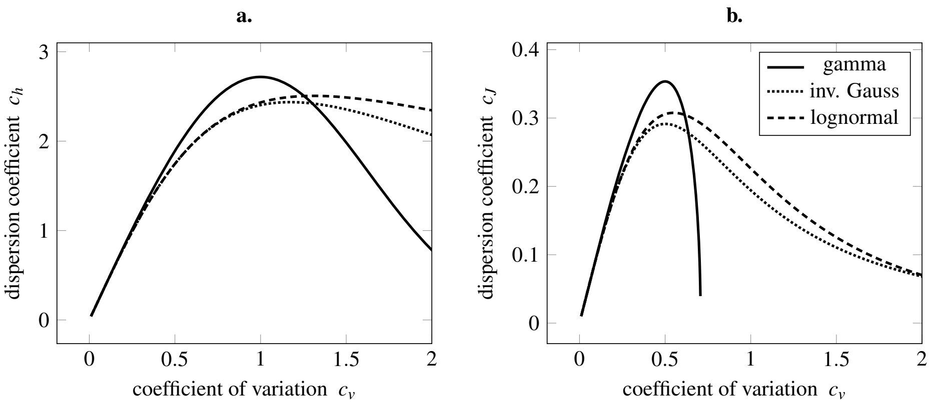

Fig. 1: Entropy-based dispersion coefficient, c_h, in dependence on the coefficient of variation, c_v, for three ISI distributions (**a**): gamma, inverse Gaussian and lognormal distribution. Both c_v and c_h describe "spread" of the interspike intervals, but from different points of view. Coefficient of variation, c_v, quantifies how off-centered is the mass of the probability density function, whereas c_h indicates how evenly is the mass distributed over all possible values. Fisher information-based dispersion coefficient, c_J, as a function of the coefficient of variation, c_v (**b**). The graph confirms that "smoothness" (given by c_J) and "evenness" (c_h) of the distribution are different notions. Still, there are qualitative similarities: $c_J = 0$ for $c_v \to 0$ for all shown distributions, and $c_J = 0$ as $c_v \to \infty$ for both lognormal and inverse Gaussian distributions.

4 Discussion and conclusions

In this contribution, we aim to point out the difference between frequently interchanged notions of variability and randomness. Variability described by c_v and randomness by c_h are different concepts. Consider, for example, a spike train consisting of "long" and "short" ISIs with no serial correlations. By adding "medium" length ISIs the spiking variability is not increased, contrary to what is expected intuitively, but is decreased. Simultaneously, since the count of ISI of different lengths increases, the spiking randomness is increased. Thus, even if a conventional analysis of two spike trains reveals no difference, the spike trains may still differ in their randomness and the difference is tractable with a limited amount of data [3].

Additionally, by considering the Fisher information-based dispersion coefficient, c_J, we show that ISI randomness (increasing with diversity of the ISI lengths) and probability density "smoothness" are related, but still different notions. For example, all of the tested distributions are "maximally smooth" for $c_v \doteq 0.5$ and "maximally even" (maximum ISI randomness) for $c_v \doteq 1$.

References

[1] J. F. Bercher and C. Vignat. On minimum Fisher information distributions with restricted support and fixed variance. *Inform. Sciences*, 179(22):3832–3842, 2009.

[2] B. R. Frieden. *Physics from Fisher information: a unification.* Cambridge University Press, New York, 1998.

[3] L. Kostal, P. Lansky, and J-P. Rospars. Review: Neuronal coding and spiking randomness. *Eur. J. Neurosci.*, 26:2693–2701, 2007.

[4] L. Kostal, P. Lansky, and O. Pokora. Variability measures of positive random variables. *PLoS ONE*, 6(7):e21998, 2011.

[5] V. Zivojnovic. A robust accuracy improvement method for blind identification usinghigher order statistics. In *IEEE Internat. Conf. Acous. Speech. Signal Proc., 1993. ICASSP-93.*, volume 4, pages 516–519, 1993.

[6] L. Kostal, P. Lansky, and C. Zucca. Randomness and variability of the neuronal activity described by the Ornstein-Uhlenbeck model. *Network: Comput. Neural Syst.*, 18:63–75, 2007.

[7] H. C. Tuckwell. *Introduction to Theoretical Neurobiology, Volume 2.* Cambridge University Press, New York, 1988.

[8] M. Kendall, A. Stuart, and J. K. Ord. *The advanced theory of statistics. Vol. 1: Distribution theory.* Charles Griffin, London, 1977.

[9] S. Ditlevsen and P. Lansky. Firing variability is higher than deduced from the empirical coefficient of variation. *Neural Comput., in press*, 2011.

[10] R. B. Ash. *Information Theory.* Dover, New York, 1965.

[11] L. Kostal and P. Marsalek. Neuronal jitter: can we measure the spike timing dispersion differently? *Chin. J. Physiol.*, 53(6):454–464, 2010.

[12] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* John Wiley and Sons, Inc., New York, 1991.

[13] E J G Pitman. *Some basic theory for statistical inference.* John Wiley and Sons, Inc., New York, 1979.

[14] P. Duchamp-Viret, L. Kostal, M. Chaput, P. Lansky, and J-P. Rospars. Patterns of spontaneous activity in single rat olfactory receptor neurons are different in normally breathing and tracheotomized animals. *J. Neurobiol.*, 65(2):97–114, 2005.

[15] C. Pouzat and A. Chaffiol. Automatic Spike Train Analysis and Report Generation. An Implementation with R, R2HTML and STAR. *J. Neurosci. Methods*, 181(1):119–144, 2009.

[16] D. Berger, K. Pribram, H. Wild, and C. Bridges. An analysis of neural spike-train distributions: determinants of the response of visual cortex neurons to changes in orientation and spatial frequency. *Exp. Brain Res.*, 80(1):129–134, 1990.

[17] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables.* Dover, New York, 1965.

[18] P. Lansky and S. Sato. The stochastic diffusion models of nerve membrane depolarization and interspike interval generation. *J. Peripher. Nerv. Syst.*, 4(1):27–42, 1999.