Neuronal Data Analysis Based on the Empirical Cumulative Entropy

Antonio Di Crescenzo[1] and Maria Longobardi[2]

[1] Dipartimento di Matematica, Università di Salerno
I-84084 Fisciano (SA), Italy
adicrescenzo@unisa.it
[2] Dipartimento di Matematica e Applicazioni, Università di Napoli Federico II
I-80126 Napoli, Italy
maria.longobardi@unina.it

Abstract. We propose the empirical cumulative entropy as a variability measure suitable to describe the information content in neuronal firing data. Some useful characteristics and an application to a real dataset are also discussed.

1 Introduction

The cumulative entropy has been proposed recently in Di Crescenzo and Longobardi [3] as a measure of the uncertainty contained in a random variable. It plays a relevant role in reliability theory, being particularly suitable to describe the information in problems related to ageing properties based on the past and on the inactivity times. The empirical cumulative entropy has been considered in Di Crescenzo and Longobardi [4] in order to estimate the information content of the cumulative entropy in random lifetimes. It has been shown to depend on the dual normalized sample spacings, and to converge to the cumulative entropy as the sample size goes to infinity.

In this paper we show some connections between the partition entropy, the cumulative entropy and the cumulative residual entropy. Moreover, we illustrate the usefulness of the empirical cumulative entropy to the analysis of variability in temporal patterns in neuronal coding. Attention is given to the measure of distance between neuronal spykes, by taking as reference the sample data presented in Kass *et al.* [9].

2 Entropies

Consider an absolutely continuous non-negative random variable X; let $F(x) = P(X \le x)$ be its distribution function and $\overline{F}(x) = 1 - F(x)$ its cumulative residual distribution (also known as survival function). Denote by $f(x)$ the probability density function of X. According to Bowden [1], the function

$$h(x) = -[F(x)\log F(x) + \overline{F}(x)\log \overline{F}(x)], \qquad x \ge 0 \tag{1}$$

R. Moreno-Díaz et al. (Eds.): EUROCAST 2011, Part I, LNCS 6927, pp. 72–79, 2012.

is the *partition entropy* at x. It can be seen as a measure of the information derived from knowing whether X takes values in $[0, x]$ or in $(x, +\infty)$. Indeed, denoting the Shannon *differential entropy* of X by

$$H := -\mathbb{E}[\log f(X)] = -\int_0^{+\infty} f(x) \log f(x) \, \mathrm{d}x,$$

for all $t > 0$ we have (cf. Proposition 2.1 of Di Crescenzo and Longobardi [2]):

$$H = h(t) + F(t)\, \overline{H}(t) + \overline{F}(t)\, H(t), \tag{2}$$

In Eq. (2) $H(t)$ and $\overline{H}(t)$ denote respectively the *residual entropy* and the *past entropy* of X, defined for $t > 0$ as

$$H(t) = -\int_t^{+\infty} \frac{f(x)}{\overline{F}(t)} \log \frac{f(x)}{\overline{F}(t)} \, \mathrm{d}x, \qquad \overline{H}(t) = -\int_0^{t} \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)} \, \mathrm{d}x.$$

These functions have been introduced by Ebrahimi and Pellerey [6] and by Di Crescenzo and Longobardi [2] in order to describe uncertainty in residual lifetime distributions and in past lifetime distributions, respectively.

More recently, aiming to provide an alternative to the differential entropy for the description of information in stochastic systems, the following new uncertainty measures have been proposed:

$$\mathcal{E}(X) = -\int_0^{+\infty} \overline{F}(x)\, \log \overline{F}(x) \, \mathrm{d}x, \qquad \mathcal{CE}(X) = -\int_0^{+\infty} F(x)\, \log F(x) \, \mathrm{d}x. \tag{3}$$

The first measure is known as *cumulative residual entropy*, and has been studied by Rao [17] and Rao *et al.* [18]. We also recall that $\mathcal{E}(X)$ has been applied to image analysis in [24]. The second measure is named *cumulative entropy*. It has been proposed and investigated by Di Crescenzo and Longobardi [3], [4] and [5].

The measures $\mathcal{E}(X)$ and $\mathcal{CE}(X)$ can be viewed as dual, since they are useful to describe information in dynamic reliability systems when uncertainty is related to the future, and to the past, respectively. Moreover, from (1) and (3) we have the following nice relation between the partition entropy, the cumulative residual entropy and the cumulative entropy:

$$\mathcal{E}(X) + \mathcal{CE}(X) = \int_0^{+\infty} h(x) \, \mathrm{d}x.$$

3 Empirical Cumulative Entropy

It has been pointed out in Di Crescenzo and Longobardi [3] that $\mathcal{CE}(X)$ can be estimated by means of the *empirical cumulative entropy*. Given a random sample $X_1, X_2, \ldots, X_n$ of non-negative random variables the empirical cumulative entropy is defined as

$$\mathcal{CE}(\hat{F}_n) = -\int_0^{+\infty} \hat{F}_n(x)\, \log \hat{F}_n(x) \, \mathrm{d}x,$$

which is expressed in terms of the empirical distribution of the sample

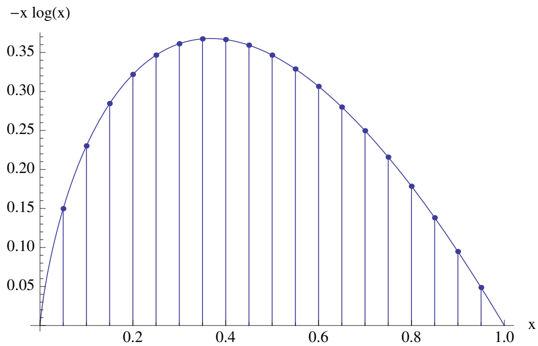

Fig. 1. Function $-x\log(x)$ with indication of weights $\frac{j}{n}\log\frac{j}{n}$ for $1 \le j \le n-1$, and $n = 20$

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \le x\}}, \qquad x \in \mathbb{R}.$$

Asymptotic results of $\mathcal{CE}(\hat{F}_n)$ involve its a.s. convergence to $\mathcal{CE}(X)$ (see Proposition 2 of [4]) and a central limit theorem in the case of exponentially distributed random variables (see Theorem 7.1 of [3]). Denoting by $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$ the order statistics of the random sample, we also notice that $\mathcal{CE}(\hat{F}_n)$ can be expressed as

$$\mathcal{CE}(\hat{F}_n) = -\sum_{j=1}^{n-1} U_{j+1}\,\frac{j}{n}\log\frac{j}{n}, \tag{4}$$

where

$$U_1 = X_{(1)}, \qquad U_i = X_{(i)} - X_{(i-1)}, \qquad i = 2, 3, \ldots, n,$$

are the sample spacings. Hence, Eq. (4) expresses that the empirical cumulative entropy is a linear combination (with positive weights) of the sample spacings $U_2, \ldots, U_n$. Since the function $-x\log(x)$ is concave and attains its maximum for $x = \mathrm{e}^{-1} \approx 0.3679$, the outer spacings U_2 and U_n possess small weights (i.e., $\frac{1}{n}\log\frac{1}{n}$ and $\frac{n-1}{n}\log\frac{n-1}{n}$), whereas the larger weight is given to the spacing U_{j+1} such that j is close to $0.3679\,n$. As example, Figure 1 shows the case $n = 20$, where the larger weight $\frac{j}{n}\log\frac{j}{n}$ corresponds to $j = 0.35$.

In other terms, the empirical cumulative entropy is asymmetric in the sense that it measures the distance between adjacent ordered sample data by giving asymmetric weights to such distances. Hence, it is particularly suitable to measure variability in data distributions that are skewed to the right, such as those describing the firing times in neuronal activity. This is confirmed in the following example, where for a suitable family of random variables the minimum of the variance is attained in the middle point of the range of the parameter θ (i.e., when the density is symmetric), whereas for the cumulative entropy the minimum is located in a different point, situated on the left (by which the density is positively skewed).

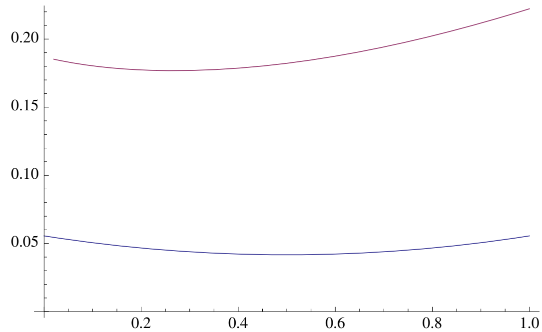

Fig. 2. Plot of $\mathcal{CE}(\hat{F}_n)$ (top) and $\mathbb{V}ar(X)$ (bottom) as $\theta \in (0,1)$ in Example 1

Example 1. Consider the following family of triangular probability density functions, where θ is the location of the mode:

$$f_X(x) = \begin{cases} \dfrac{2x}{\theta}, & 0 \le x \le \theta \\ \dfrac{2(x-1)}{\theta - 1}, & \theta \le x \le 1 \end{cases} \qquad (0 \le \theta \le 1).$$

Since $\mathbb{V}ar(X) = (1 - \theta + \theta^2)/18$, it attains its minimum for $\theta = 0.5$, whereas the cumulative entropy is minimum for $\theta \approx 0.27$, as shown in Figure 2.

4 Neural Coding

In this section we discuss the application of the empirical cumulative entropy to neuronal firing data.

The relevance of computational methods of information theory in studying and understanding the neural coding has been pinpointed in many investigations. See, for instance, the reviews by Kostál and Lánský [12] and Kostál *et al.* [13] on some recently proposed measures of randomness compared to the coefficient of variation, which is the frequently employed measure of variability of spiking neuronal activity.

Among several contributions we recall the paper by Kostál and Lánský [10], which deals with a measure of the information rate of a single stationary neuronal activity; such measure is based on the Kullback-Leibler (KL) distance between two interspike interval distributions. The KL information distance has also been used in the field of neuronal coding as a tool in classification problems (cf. Johnson *et al.* [8]), as a predictor of purely rate coding models (cf. Johnson and Glantz [7]), and for judging similarity between two different interspike interval distributions (cf. Kostál and Lánský [11]).

Another kind of information measure, suitably defined in terms of the differential entropy, is the so-called normalized entropy. This is explored in Kostál *et al.* [15] as a measure of randomness of neuronal firings concerning the Ornstein-Uhlenbeck neuronal model (see Ricciardi [19], Ricciardi and Sacerdote [20] and

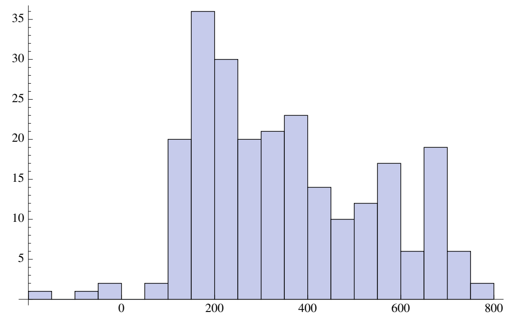

Fig. 3. Histogram of 242 neural sample data by Kass *et al.* [9]

Ricciardi *et al.* [21] for comprehensive descriptions of such stochastic model). A detailed discussion of the properties of the normalized entropy, and comparisons with the widely employed standard deviation measure are given in Kostál and Marsalek [14].

Moreover, on the ground of a distance measure proposed by Victor and Purpura [23] and based on a suitable cost function, a sort of related discrimination measure for the distance between two spike trains is proposed by Van Rossum [22] to measure the intrinsic noise of a model neuron.

In this framework we propose to adopt the empirical cumulative entropy as a new suitable tool able to ascertain information in the context of neuronal firing data. Indeed, as suggested in Example 1 for the cumulative entropy, this novel measure is especially suitable to fit skewed-to-the-right distributions, such as those typically involved in neural activity. Hereafter we employ the empirical cumulative entropy in a preliminary example, aiming to discuss its effective characteristics in a subsequent paper.

Example 2. We consider a case-study based on a dataset of 242 spike times observed in 8 trials on a single neuron. Data are taken from the paper by Kass *et al.* [9]. The corresponding histogram is shown in Figure 3.
The spike times of the first trial are listed hereafter ($n = 29$):

$$\{136.842, 145.965, 155.088, 175.439, 184.561, 199.298, 221.053, 231.579, 246.316, 263.158, 274.386, 282.105, 317.193, 329.123, 347.368, 360.702, 368.421, 389.474, 392.982, 432.281, 449.123, 463.86, 503.86, 538.947, 586.667, 596.491, 658.246, 668.772, 684.912\}.$$

These data lead to the following estimated cumulative entropy:

$$\mathcal{CE}(\hat{F}_{29}) = -\sum_{j=1}^{n-1} U_{j+1} \frac{j}{n} \log \frac{j}{n} = 131.223.$$

For each of the 8 trials we exhibit the empirical cumulative entropy in Figure 4, whereas the mean firing times and the standard deviation are shown in Figure 5. We notice that the orderings among the empirical cumulative entropies, the

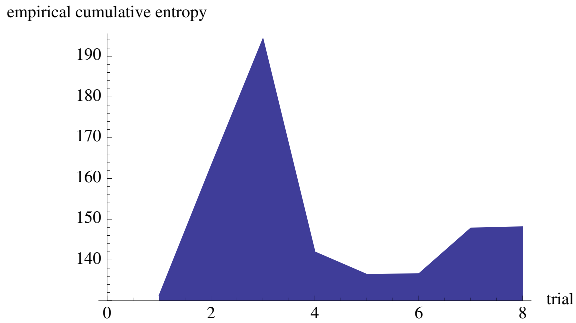

Fig. 4. The empirical cumulative entropy of the 8 trials

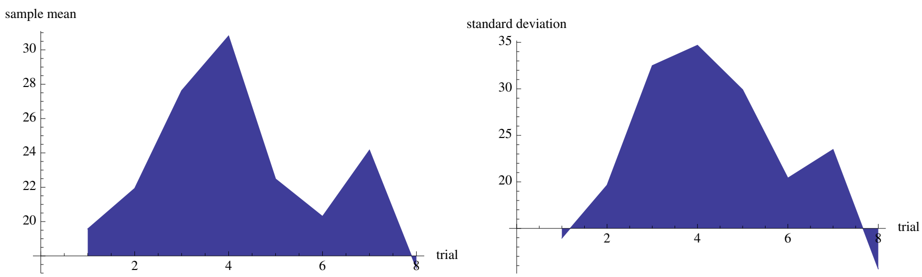

Fig. 5. The mean firing time and the standard deviation of the 8 trials

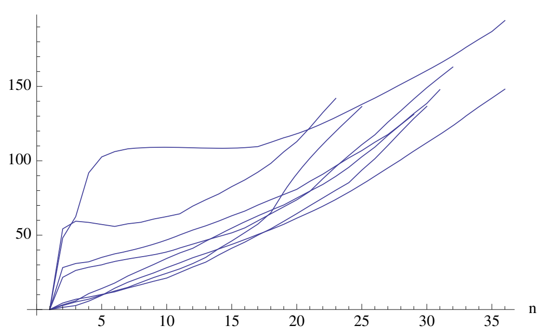

Fig. 6. The empirical cumulative entropy of the first n data of the 8 trials

sample means and the standard deviations of the 8 trials are not identical. This show that the empirical cumulative entropies allows to gain information on the variability in the sample data quite different from that provided by classical statistical indices.

We finally remark that, in general, if the sample data are arranged in increasing order, then the empirical cumulative entropy is not necessarily increasing in n (see [5]). This is confirmed by Figure 6, where for each dataset of the 8 trials we plot the empirical cumulative entropy of the first n data, by assuming that the sample data are observed in increasing order. For certain trials the resulting curve are shown to be not increasing.

5 Concluding Remarks

In conclusion we pinpoint that one of the crucial and actual problems of neuroscientists is to understand the neural coding adopted by the brain to handle information. Up to now the following 2 hypothesis were considered: *(i)* information is based on the neuronal spike frequency (the number of spikes in the time unit); *(ii)* information is based on the temporal occurrence of spikes (the sequence of spikes). In the past a common working assumption stated that it is impossible that the two above neural coding coexist. On the contrary, recent wide analysis on current experimental results suggest that the two codes may be adopted simultaneously by the brain (see Kumar *et al.* [16]).

On the ground of these novel findings, in a future research we plan to generalize the cumulative entropy to the two-dimensional case, in order to use this new measure to discriminate and describe information in multivariate patterns based on both the number and the sequence of neural spikes.

Acknowledgments. This work has been partially supported by Ministero dell'Istruzione, dell'Università e della Ricerca within the Project "Mathematical models and computation methods for information processing and transmission in neuronal systems subject to stochastic dynamics" (PRIN 2008), directed by Professor Luigi M. Ricciardi (from University of Naples Federico II), who sadly passed away while this paper was about to be completed. The authors would like to dedicate this article and their work within this project to his memory.

References

1. Bowden R.J., Information, measure shifts and distribution metrics, Statistics, First published on (September 2, 2010)
2. Di Crescenzo, A., Longobardi, M.: Entropy-based measure of uncertainty in past lifetime distributions. J. Appl. Prob. 39, 434–440 (2002)
3. Di Crescenzo, A., Longobardi, M.: On cumulative entropies. J. Stat. Plann. Infer. 139, 4072–4087 (2009)
4. Di Crescenzo, A., Longobardi, M.: On cumulative entropies and lifetime estimations. In: Mira, J., Ferrández, J.M., Álvarez, J.R., de la Paz, F., Toledo, F.J. (eds.) IWINAC 2009 Part I. LNCS, vol. 5601, pp. 132–141. Springer, Heidelberg (2009)
5. Di Crescenzo, A., Longobardi, M.: More on cumulative entropy. In: Trappl, R. (ed.) Cybernetics and Systems 2010, pp. 181–186. Austrian Society for Cybernetic Studies, Vienna (2010)
6. Ebrahimi, N., Pellerey, F.: New partial ordering of survival functions based on the notion of uncertainty. J. Appl. Prob. 32, 202–211 (1995)
7. Johnson, D., Glantz, R.: When does interval coding occur? Neurocomputing 59, 13–18 (2004)
8. Johnson, D., Gruner, C., Baggerly, K., Seshagiri, C.: Information theoretic analysis of neural coding. J. Comp. Neurosci. 10, 47–69 (2001)

9. Kass, R.E., Ventura, V., Cai, C.: Statistical smoothing of neuronal data. Network: Comput. Neur. Syst. 14, 5–15 (2003)
10. Kostál, L., Lánský, P.: Classification of stationary neuronal activity according to its information rate. Network: Comput. Neur. Syst. 17, 193–210 (2006)
11. Kostál, L., Lánský, P.: Similarity of interspike interval distributions and information gain in a stationary neuronal firing. Biol. Cybern. 94, 157–167 (2006)
12. Kostál, L., Lánský, P.: Randomness of spontaneous activity and information transfer in neurons. Physiol. Res. 57, S133-S138 (2008)
13. Kostál, L., Lánský, P., Rospars, J.P.: Neuronal coding and spiking randomness. European J. Neurosci. 26, 2693–2701 (2007)
14. Kostál, L., Marsalek, P.: Neuronal jitter: can we measure the spike timing dispersion differently? Chinese J. Physiol. 53, 454–464 (2010)
15. Kostál, L., Lánský, P., Zucca, C.: Randomness and variability of the neuronal activity described by the Ornstein–Uhlenbeck model. Network: Comput. Neur. Syst. 18, 63–75 (2007)
16. Kumar, A., Rotter, S., Aertsen, A.: Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. Nature Rev. Neurosci. 11, 615–627 (2010)
17. Rao, M.: More on a new concept of entropy and information. J. Theor. Probab. 18, 967–981 (2005)
18. Rao, M., Chen, Y., Vemuri, B.C., Wang, F.: Cumulative residual entropy: a new measure of information. IEEE Trans. Inform. Theory 50, 1220–1228 (2004)
19. Ricciardi, L.M.: Diffusion processes and related topics in biology. Springer, Berlin (1977)
20. Ricciardi, L.M., Sacerdote, L.: The Ornstein-Uhlenbeck process as a model for neuronal activity I. Mean and variance of the firing time. Biol. Cybern. 35, 1–9 (1979)
21. Ricciardi, L.M., Di Crescenzo, A., Giorno, V., Nobile, A.G.: An outline of theoretical and algorithmic approaches to first passage time problems with applications to biological modeling. Math. Japon. 50, 247–322 (1999)
22. Van Rossum, M.C.W.: A novel spike distance. Neural Comput. 13, 751–763 (2001)
23. Victor, J.D., Purpura, K.P.: Metric-space analysis of spike trains: theory, algorithms and application. Network: Comput. Neur. Syst. 8, 127–164 (1997)
24. Wang, F., Vemuri, B.C.: Non-rigid multi-modal image registration using cross-cumulative residual entropy. Intern. J. Comp. Vision 74, 201–215 (2007)