Physica A 390 (2011) 154–160

Contents lists available at ScienceDirect

Physica A

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

General model of subtraction of stochastic variables. Attractor and stability analysis

M. Beltrán del Río [a,*], G. Cocho [a,b], R. Mansilla [b,c]

[a] Departamento de Sistemas Complejos, Instituto de Física, UNAM, 04510, D. F., Mexico

[b] Centro de Ciencias de la Complejidad, UNAM, 04510, D. F, Mexico

[c] Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades, UNAM, 04510, D. F., Mexico

ARTICLE INFO

Article history:
Received 12 April 2010
Received in revised form 24 September 2010
Available online 7 October 2010

Keywords:
Central Limit Theorem
Random systems

ABSTRACT

We introduce a general process designed to model stochastic systems in which the dependence of random variables is not through addition only but combined addition and subtraction with bounded ranges, and whose probabilistic factors have compact support. We show that, still retaining much of the general essence of the Central Limit Theorem, this process presents a functional attractor which is neither Gaussian nor Lévy like, and is precisely akin numerically to a probability density function shown in previous works to have ubiquitous character, namely the two-parameter beta distribution.

1. Introduction

The search for regularities has been one of the leading forces in the attempts to understand natural phenomena. The story of power laws dates back a long time [1] and since then power laws have been used to describe many different sets of empirical data. Nevertheless it is seldom seen that the fit of real data to the theoretical power law shows two breakdowns for small and large values of the independent variable. Several explanations have been provided for this phenomenon, standing out among them is the so-called finite size effect (e.g. not enough data for good statistics) [2,3]. Recently, a new ansatz has been proposed [4,5] yielding a notably good fit to empirical data in the whole range of the independent variable. In this paper some properties of this function are studied. The connection between these properties and some fundamental results of probability theory are also established. The structure of the paper is as follows: In Section 2 we state the main idea of the new model, outline the general settings of the Central Limit Theorem (CLT from now on) and its relationship with our results. In Section 3 we discuss how the Two-Parameter Beta Distribution (TPBD from now on) is stable under the contrived integral operator just as Gaussian and Lévy distributions are under convolution, drawing a parallel with the CLT. Section 4 presents concluding remarks.

2. Addition vs. subtraction of random variables

Groups of different systems that depend on a large number of different and independent events or influences can show regular statistical properties, common to a wide variety of dynamics. A typical example of this is how the widespread presence of the Normal Distribution in diverse situations involved with a multiplicity of factors is frequently explained by the CLT. Simply put, this theorem states that a variable which depends additively on many random variables will be well fitted around the mean by a Normal Distribution. The greater the number of these random factors the better the fit will

* Corresponding author.
E-mail addresses: manubrg@gmail.com (M. Beltrán del Río), cocho@fisica.unam.mx (G. Cocho).

0378-4371/$ – see front matter
doi:10.1016/j.physa.2010.09.035

be. Our current interest is to focus on systems whose output depends on the bounded difference of stochastic variables. These systems would no longer account for the hypotheses of the CLT, specifically those that require the variables to be independent and similarly distributed. Examples of applicability are numerous given the simplicity of the model. Some single neuron models, for example, consist of the stochastic adding and subtracting of signals. It has also been shown that the neutralist species number dynamics model in Ref. [6], through a simple "birth–death" process, produces probability density functions (PDFs from now on) that agree with those in our model. The main process in both cases is the loss of information that arises from the decoherent contribution of the factors and their stochastic nature. It is therefore often fruitful when having found a strikingly ubiquitous statistical regularity, as is our case with the TPBD, to first assume that it is not due to common underlying microdynamics. Rather one should think that at some stage of the process many factors act in an uncooperative fashion leading to a certain type of information loss.

2.1. The relationship with the Central Limit Theorem

In their seminal work [7] Gnedenko and Kolmogorov studied the limit distributions of sums of independent random variables. They established that the necessary and sufficient condition for a function $F(z)$ to be the limit PDF of sums of independent and identically distributed random variables is to be *stable*. The last definition means that for every a_1, b_1, a_2, and b_2, with $a_1, a_2 > 0$, there exist $a > 0$, b such that:

$$F(a_1x + b_1) * F(a_2x + b2) = F(ax + b) \tag{1}$$

where:

$$(f * g)(z) = \int_{\Re} f(z - y)g(y)\mathbf{d}y. \tag{2}$$

According to Ref. [7, p. 162], a set of distributions is stable *if it contains all possible convolutions of elements belonging to it*. Further, the stability of a distribution is related to some properties of the characteristic function of the distribution.

There are however no results similar to the aforementioned if we drop the assumption of independence among the elements of the sum of random variables.

2.2. Difference distribution product

The probability density function of an additive process depending on multiple, independent, stochastic variables is obtained naturally through the reiterative application of convolution:

$$P_2 \oplus P_1(\chi) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{d}\xi_1\mathbf{d}\xi_2 P_1(\xi_1)P_2(\xi_2)\delta(\xi_2 + \xi_1 - \chi). \tag{3}$$

In convolution the integration is done, as the δ function expresses, in the locus of equal sum of variables. If we are interested in the probabilistic outcome of the variable $\chi = \chi_2 - \chi_1$, each following PDFs $P_1(\chi_2)$ and $P_1(\chi_1)$, then:

$$\begin{aligned} P(\chi) &\equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{d}\xi_1\mathbf{d}\xi_2 P_1(\xi_1)P_2(\xi_2)\delta(\xi_2 - \xi_1 - \chi) \\ &= P_2(\chi_2) \oplus P_1(-\chi_1). \end{aligned} \tag{4}$$

Sequential application of this operation would in general no longer account for the hypotheses of the CLT, in particular because it might introduce a change in the expected value of the PDF $P_1(-\chi_1)$ and thus possibly an anomalous drift of the resulting $P(\chi)$. Introducing this change we have devised a process in which the stochastic outcome is governed by alternating sums and subtractions of stochastic variables in a particular order, in contrast with the simple unordered sums of the CLT. We further demand finite positive support for the PDFs and also we only perform the integration for positive values of the variable χ, so discarding all probability of a negative outcome. Thus, $\chi_i \in (0, 1)$ and $\chi = \chi_2 - \chi_1 < 0$ is not counted in the probability integral. If we add these conditions we can then define a new product:

$$\begin{aligned} (P_1 \ominus P_2)(\chi) &\equiv N\int_0^1\int_0^1 \mathbf{d}\xi_1\mathbf{d}\xi_2 P_1(\xi_1)P_2(\xi_2)\delta(\xi_2 - \xi_1 - \chi)\Theta(\chi) \\ &= N\int_{\chi}^1 \mathbf{d}\xi P_1(\xi - \chi)P_2(\xi)\Theta(\chi), \end{aligned} \tag{5}$$

which is a cross-correlation integral but with the sign of the variable changed and restricted by a Heaviside Θ function of the argument. Notice that this product is not abelian and that it needs a normalisation factor N, unlike convolution. Convolution guarantees normalisation without the need of introducing any factor if it acts on two PDFs, since

$$P = P_2 \oplus P_1 \Rightarrow \hat{P} = \hat{P_1}\hat{P_2},$$