Physica A 515 (2019) 248–257

Contents lists available at ScienceDirect

Physica A

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

Use of the geometric mean as a statistic for the scale of the coupled Gaussian distributions

Kenric P. Nelson [a,*], Mark A. Kon [a], Sabir R. Umarov [b]

[a] Boston University, United States
[b] University of New Haven, United States

HIGHLIGHTS

- Density of coupled Gaussian at mean plus scale is geometric mean of densities.
- Estimator of scale using geometric mean of random samples established.
- Numerical evidence suggests this estimator is unbiased and has diminishing variance.
- Fluctuations shown to be proportional to the nonlinear statistical coupling.

ARTICLE INFO

Article history:
Received 5 March 2018
Received in revised form 11 August 2018
Available online xxxx

Keywords:
Complex systems
Information theory
Nonextensive statistical mechanics
Heavy-tail

ABSTRACT

The geometric mean is shown to be an appropriate statistic for the scale of a heavy-tailed coupled Gaussian distribution or equivalently the Student's t distribution. The coupled Gaussian is a member of a family of distributions parameterized by the nonlinear statistical coupling which is the reciprocal of the degree of freedom and is proportional to fluctuations in the inverse scale of the Gaussian. Existing estimators of the scale of the coupled Gaussian have relied on estimates of the full distribution, and they suffer from problems related to outliers in heavy-tailed distributions. In this paper, the scale of a coupled Gaussian is proven to be equal to the product of the generalized mean and the square root of the coupling. From our numerical computations of the scales of coupled Gaussians using the generalized mean of random samples, it is indicated that only samples from a Cauchy distribution (with coupling parameter one) form an unbiased estimate with diminishing variance for large samples. Nevertheless, we also prove that the scale is a function of the geometric mean, the coupling term and a harmonic number. Numerical experiments show that this estimator is unbiased with diminishing variance for large samples for a broad range of coupling values.

1. Introduction

The estimation of uncertainty rests principally on the variance and entropy of a random variable, but in nonlinear systems a broader perspective on statistical analysis is required. One manifestation of the complexity of nonlinear systems is the role of power law statistics, which are not characterized by the variance of related random variables. Renyi, Tsallis and other investigators have shown that generalizations of the entropy function can be used to explain the role of power law statistics in complex systems [1–4]. It is not as widely appreciated, that both the Renyi and Tsallis entropies of a distribution utilize the generalized mean to combine the probabilities of a distribution. The Rényi entropy translates the generalized mean of the

* Corresponding author.
E-mail address: kenricpn@bu.edu (K.P. Nelson).

https://doi.org/10.1016/j.physa.2018.09.049

distribution to an entropy scale using the natural logarithm, while the Tsallis entropy utilizes a generalized logarithm [5,6]. The special case of Boltzmann–Gibbs–Shannon entropy is a translation of the geometric mean by the logarithm. As such the generalized mean is an important statistic for characterizing uncertainty in data arising from complex systems. In this paper we explore the properties of the generalized mean as an estimator of the scale for a family of distributions which have power-law decay.

Two fundamental criteria for estimation are efficiency [7] and sufficiency [8]. An estimator is said to be efficient if it minimizes a loss function, such as the mean-square error. The mean square error of an estimator $\hat{\theta}$ is the sum of the variance and the square of the bias

$$MSE = E\left[\left(\hat{\theta}-\theta\right)^2\right] = E\left[\left(\hat{\theta}-E\left(\hat{\theta}\right)+E\left(\hat{\theta}\right)-\theta\right)^2\right] = Var\left(\hat{\theta}\right)+\left(b\left(\hat{\theta}\right)\right)^2. \tag{1}$$

An estimator which is unbiased and minimizes the variance, would thus be an efficient unbiased estimator based on the mean-square error criteria. A stronger criterion is sufficiency, which requires that a statistic utilize all the possible information in a dataset with regard to estimating a parameter. More precisely, sufficiency requires that the conditional probability given a statistic is independent of the parameter

$$p\left(\mathbf{X}=\mathbf{x}|T\left(\mathbf{X}\right)=t,\theta\right)=p\left(\mathbf{X}=\mathbf{x}|T\left(\mathbf{X}\right)=t\right). \tag{2}$$

In other words, given the statistic t of the dataset $\mathbf{x}$, knowledge of the parameter does not reduce the uncertainty. A statistic is sufficient if and only if the probability of the dataset conditioned on the parameter can be factored into two functions, one dependent on the statistic and one independent of the statistic,

$$f\left(\mathbf{x}|\theta\right)=g\left(T\left(\mathbf{x}\right)|\theta\right)h\left(\mathbf{x}\right). \tag{3}$$

The exponential family is defined in terms of this factorization principal [9]. Given the natural parameters $\eta(\theta)$, their sufficient statistics $T(x)$ and the log partition function $A(\theta)$ the exponential family of distributions is defined as

$$f\left(x|\theta\right)=h\left(x\right)\exp\left(\eta\left(\theta\right)T\left(x\right)-A\left(\theta\right)\right). \tag{4}$$

In this paper, we provide numerical evidence that a family of heavy-tail distributions referred to as coupled-exponentials [10], may have an efficient estimator for their scale parameter. If so, there may be a generalization of the factorization theorem for sufficiency. The coupled exponentials are non-exponential except in the degenerate case that the coupling is zero. The coupling refers to a nonlinear deformation of the exponential and logarithmic functions. This generalization is reviewed in Section 2. In Section 3, numerical experiments using the generalized mean as a statistic to estimate the scale parameter of the coupled exponential distributions are explored. Evidence is provided that the geometric mean in particular has potential to be a statistic from which an unbiased efficient estimator of the scale can be formed. Section 4 provides examples of estimating the scale and tail decay of the coupled exponential distributions. Section 5 provides a conclusion and discussion of the results.

2. Review of the coupled exponential family of distributions

The investigation focuses on a family of distributions [11] related to the power function $(1+\kappa x)^{\frac{1}{\kappa}}$, which reduces to the exponential function as $\kappa \to 0$. For $\kappa \neq 0$ this function approximates the exponential near $x=0$ and has power-law properties as $x\to\infty$. This function solves the nonlinear differential equation $\frac{dy}{dx}=y^{1-\kappa}$. For this and further reasons (explained below) κ is referred to as the *nonlinear statistical coupling* [12] or simply the *coupling*. We use a shorthand notation for the power function emphasizing its role as a deformation of the exponential function [13] by writing

$$\exp_\kappa^a\left(x\right)\equiv\left(1+\kappa x\right)_+^{\frac{a}{\kappa}}, \tag{5}$$

where the parameter measures the level of deformation from the exponential function, while the exponent a is simply a power parameter; here we have defined $(y)_+=\max(0,y)$. Given an argument $\left|\frac{x-\mu}{\sigma}\right|^\alpha$ with mean μ, scale σ and power α, a family of distributions is defined which we will refer to as the *coupled exponential family*.

Definition 1. A coupled exponential family of distributions is defined by

$$\begin{gathered} f\left(x;\mu,\sigma,\kappa,\alpha\right)\equiv\frac{1}{Z\left(\sigma,\kappa,\alpha\right)}\exp_\kappa^{-\frac{1+\kappa}{\alpha}}\left(\frac{|x-\mu|^\alpha}{\sigma^\alpha}\right)=\left(Z\left(\sigma,\kappa,\alpha\right)\left(1+\kappa\frac{|x-\mu|^\alpha}{\sigma^\alpha}\right)_+^{\frac{1+\kappa}{\alpha\kappa}}\right)^{-1}, \\ Z\left(\sigma,\kappa,\alpha\right)=\begin{cases}\dfrac{2\sigma}{\alpha}\kappa^{-\frac{1}{\alpha}}\mathrm{B}\left(\dfrac{1}{\alpha\kappa},\dfrac{1}{\alpha}\right), & \kappa>0\\ \dfrac{2\sigma}{\Gamma\left(1+\frac{1}{\alpha}\right)}, & \kappa=0\\ \dfrac{2\sigma}{\alpha}(-\kappa)^{-\frac{1}{\alpha}}\mathrm{B}\left(1-\dfrac{1+\kappa}{\alpha\kappa},\dfrac{1}{\alpha}\right) & -1<\kappa<0\end{cases} \\ \sigma\geq 0, 0<\alpha\leq 3, \end{gathered} \tag{6}$$