Statistics and Probability Letters 118 (2016) 156–162

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Central Limit Theorems under additive deformations

Daniel J. Eck [a,*], Ian W. McKeague [b]

[a] Department of Statistics, University of Minnesota, 313 Ford Hall, 224 Church St SE, Minneapolis, MN 55455, USA

[b] Department of Biostatistics, Columbia University, 722 West 168th St. NY, NY 10032, USA

ARTICLE INFO

Article history:
Received 4 April 2016
Received in revised form 7 June 2016
Accepted 10 June 2016
Available online 28 June 2016

Keywords:
Probability on Lie groups
Kaniadakis addition
Tsallis entropy

ABSTRACT

Additive deformations of statistical systems arise in various areas of physics. Classical central limit theory is then no longer applicable, even when standard independence assumptions are preserved. This paper investigates ways in which deformed algebraic operations lead to distinctive central limit theory. We establish some general central limit results that are applicable to a range of examples arising in nonextensive statistical mechanics, including the addition of momenta and velocities via Kaniadakis addition, and Tsallis addition. We also investigate extensions to random additive deformations, and find evidence (based on simulation studies) for a universal limit specific to each statistical system.

1. Introduction

The classical central limit theorem (CLT) is a cornerstone of statistics. In this article, we generalize this classical result to settings in which standard addition on the real line is replaced by a binary operation that satisfies Lie group properties. Additional mild smoothness assumptions are also imposed, allowing us to obtain explicit limiting distributions.

Our principal motivation comes from physics. As explained by Tempesta (2011), different Lie group operations on the real line are associated with distinctive forms of entropy that extend Boltzmann–Gibbs entropy, which corresponds to standard addition and classical central limit theory. Tsallis entropy applies to statistical systems exhibiting the features of long range dependence (Tsallis, 1988), and has been successfully applied, for example, in image thresholding (Portes de Albuquerque et al., 2004), modeling debris flow (Singh and Cui, 2015), analyzing electromagnetic pre-seismic emissions (Potirakis et al., 2012), and modeling the distribution of momenta of cold atoms in optical lattices (Douglas et al., 2006). Kaniadakis entropy arises when combining momenta in special relativity (Kaniadakis, 2006, 2013), and its associated central limit theory has recently been developed by McKeague (2015), who showed that the limiting distributions take the form of hyperbolic functions of standard normals.

There is a general formulation of the CLT on locally compact Lie groups due to Wehn (1962), but conditions are placed on the random elements after they are logarithmically mapped into the Lie algebra (tangent space at the identity). The limit distribution is described in terms of the infinitesimal generator of a semi-group of probability measures on the Lie group, but in general it does not have an explicit form. In our setting of Lie groups on the real line, however, we are able to provide an explicit CLT using only classical conditions on the random summands and a mild smoothness condition on the associated logarithmic map. Our main result generalizes the classical CLT to this setting, and addresses an open problem raised by Tempesta (2011, Section VIII) as to whether under suitable conditions an analogue of the CLT holds for "universality classes" related to generalized types of entropy, including those mentioned above.

* Corresponding author.
E-mail addresses: eckxx049@umn.edu (D.J. Eck), im2131@cumc.columbia.edu (I.W. McKeague).

http://dx.doi.org/10.1016/j.spl.2016.06.010
0167-7152/

We also establish an extension of our main result to more severe deformations that arise when the smoothness condition on the logarithmic map is relaxed (for which a slower than $\sqrt{n}$-normalization is required). We then discuss in detail all the Lie group examples mentioned above, as well as the operation for combining velocities in special relativity, and more severe additive deformations defined via exponentiation.

Both the Tsallis and Kaniadakis universality classes involve fitting parameters, so the question naturally arises as to the effect of a random specification of such parameters on the central limit behavior of the system. We investigate this question by Monte Carlo simulation studies, and reach the somewhat surprising conclusion that there is a universal limit law in the sense that it is determined solely by the form of the deformation and the expected value of the fitting parameter.

2. CLTs under additive deformations

Our results extend the classical CLT on the real line to allow additive deformations of the following form. Standard addition is replaced by a group operation $\oplus$ defined on an open and possibly infinite interval G, with $(G, \oplus)$ assumed to be a Lie group under the usual topology on the real line. Since all Lie groups on the real line are isomorphic to their Lie algebra $(\mathbb{R}, +)$, there exists an isomorphism $g : G \to \mathbb{R}$ (that is unique up to scalar multiples) such that

$$g(x \oplus y) = g(x) + g(y) \tag{1}$$

for all $x, y \in G$. In Lie group terminology, g is called the "logarithmic" map, and its inverse $f = g^{-1}$ the "exponential" map. Let $e \in G$ be the identity, and denote $G_e = G - e$. We now give our main result showing that if g has a second order Taylor expansion around e, in which the leading term is linear, then the CLT extends to $\oplus$-addition.

Theorem 1. *Let $\{X_i\}$ be a sequence of i.i.d. G_e-valued mean-zero random variables with finite variance σ^2, and let $X_{n,i} = e + X_i/\sqrt{n}$. Suppose there exists a function $\rho : G_e \to \mathbb{R}^+$ such that*

$$\rho(x) \to 0 \quad \text{as } x \to 0, \qquad \rho(x/s) \le M \quad \text{for } x \in G_e,\ s \ge s_0 \tag{2}$$

$$|g(e + x) - x - ax^2| \le x^2 \rho(x) \quad \text{for } x \in G_e, \tag{3}$$

where a, $s_0 > 1$ and $M > 0$ are prespecified constants. Also suppose there exist constants c_1, c_2, c_3, and $s_1 > 0$, such that for all $x \in G_e$ and $s \ge s_1$,

$$s|g(e + x/s)| \le c_1 |x| 1(|x| \ge c_2) + c_3. \tag{4}$$

Then

$$X_{n,1} \oplus X_{n,2} \oplus \cdots \oplus X_{n,n} \xrightarrow{\mathcal{D}} f(Z) \tag{5}$$

where $Z \sim N(a\sigma^2, \sigma^2)$.

Remarks. 1. The key smoothness condition (3) in Theorem 1 is that g has a parabolic local approximation at the identity e. The parabola can take the general form $x \mapsto a(x - e)^2 + b(x - e)$, the only requirements being that it go through $(e, 0)$, since $g(e) = 0$, and that $b \neq 0$ (so the leading term is linear). The coefficients a and b, along with σ^2, determine the "bias" of the normal r.v. Z that appears in the limit; for simplicity we stated the result just for the case $b = 1$ (giving bias $a\sigma^2$), but the result extends to the general case, where the limit is $f(bZ_b)$ with $Z_b \sim N(a\sigma^2/b, \sigma^2)$. This follows from Theorem 1 with a changed to a/b, and the maps g and f changed to $x \mapsto g(x)/b$ and $x \mapsto f(bx)$, respectively. When g is locally approximated by a straight line $x \mapsto b(x - e)$ (i.e., $a = 0$), there is no bias.

2. In Section 3 we will examine various examples in which we can find the logarithmic map g, along with its local parabolic approximation, leading to an explicit limit distribution. A classical and well-known instance arises in connection with the CLT for products of positive r.v.s, in which case $G = (0, \infty)$, $x \oplus y = xy$ for $x, y \in G$, $e = 1$, $g = \log$, $f = \exp$, and the limit distribution is log-normal. Specifically, our result gives $\prod_{i=1}^n X_{n,i} \xrightarrow{\mathcal{D}} \exp(Z)$, where $Z \sim N(-\sigma^2/2, \sigma^2)$, where $X_i > -1$ is assumed to have mean zero and finite variance σ^2. Condition (3) holds in this case by a Taylor series expansion of $x \mapsto \log(1 + x)$ around $0 \in G_e = (-1, \infty)$, namely

$$|\log(1 + x) - x + x^2/2| \le x^2 \rho(x), \quad x > -1, \tag{6}$$

where $\rho(x) = |x/(1 + x)|$ satisfies (2) with $M = 1/(s_0 - 1)$ for any $s_0 > 1$. This expansion is verified in Section 3.2.

3. Condition (4) was only used in the proof to allow dominated convergence arguments to be applied to $\sqrt{n}g(X_1/\sqrt{n})$ and $ng(X_1/\sqrt{n})^2$. However, if X_1 is assumed to have a finite *fourth* moment then (4) is not needed and the theorem continues to hold, as shown in Lemma 1 in the Supplementary Materials (see Appendix A).