Statistics and Probability Letters 86 (2014) 61–67

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

On the central limit theorem for modulus trimmed sums

Alina Bazarova [a], István Berkes [a,*], Lajos Horváth [b]

[a] Graz University of Technology, Institute of Statistics, Kopernikusgasse 24, 8010 Graz, Austria

[b] University of Utah, Department of Mathematics, Salt Lake City, UT 84112-0090, USA

ARTICLE INFO

Article history:
Received 24 August 2013
Received in revised form 1 December 2013
Accepted 4 December 2013
Available online 22 December 2013

Keywords:
Modulus trimming
Stable distribution
Iid sums
Central limit theorem

ABSTRACT

We prove a functional central limit theorem for modulus trimmed i.i.d. variables in the domain of attraction of a nonnormal stable law. In contrast to the corresponding result under ordinary trimming, our CLT contains a random centering factor which is inevitable in the nonsymmetric case. The proof is based on the weak convergence of a two-parameter process where one of the parameters is time and the second one is the fraction of truncation.

1. Introduction

Let $X_1, X_2, \ldots$ be independent, identically distributed random variables in the domain of attraction of a stable law G with parameter $0 < \alpha < 2$. That is, assume that the partial sums $S_n = \sum_{k=1}^{n} X_k$ satisfy

$$(S_n - b_n)/a_n \xrightarrow{d} G \tag{1.1}$$

with suitable norming and centering sequences $\{a_n\}$, $\{b_n\}$. The necessary and sufficient condition for (1.1) is that F, the distribution function of X_1, satisfies

$$1 - F(x) + F(-x) = x^{-\alpha} L(x), \quad x > 0 \tag{1.2}$$

and

$$\frac{1 - F(x)}{1 - F(x) + F(-x)} \to p, \qquad \frac{F(-x)}{1 - F(x) + F(-x)} \to q \quad (x \to \infty) \tag{1.3}$$

where L is a function slowly varying at ∞ and $p, q \geq 0$, $p + q = 1$. (See e.g. Feller (1971).) In contrast to the case of finite variances, the contribution of extremal terms in the partial sums S_n is not negligible and dropping a single term can change the asymptotic behavior of the sum. Let $X_{n,1} \leq X_{n,2} \leq \cdots \leq X_{n,n}$ be the order statistics of $(X_1, X_2, \ldots, X_n)$ and put for $d \geq 1$

$$S_n^{(d)} = \sum_{j=d+1}^{n-d} X_{n,j}. \tag{1.4}$$

* Corresponding author.
E-mail addresses: bazarova@tugraz.at (A. Bazarova), berkes@tugraz.at (I. Berkes), horvath@math.utah.edu (L. Horváth).

http://dx.doi.org/10.1016/j.spl.2013.12.006

For fixed d, Le Page et al. (1981) determined the asymptotic distribution of the trimmed sum $S_n^{(d)}$ and Csörgő et al. (1986b) proved that under

$$d_n \to \infty, \quad d_n/n \to 0 \tag{1.5}$$

the trimmed sum $S_n^{(d_n)}$, suitably centered and normalized, is asymptotically normal. These results give a remarkable picture on the partial sum behavior of i.i.d. sequences in the domain of attraction of a non-normal stable law. They show that the contribution of d_n extremal terms under (1.5) already gives the stable limit distribution of the total partial sum S_n and the contribution of the remaining elements will be an asymptotically normal variable with magnitude negligible compared with S_n.

The previous results describe the effects of the extremal elements of an i.i.d. sample on their partial sum. Note, however, that other kinds of trimming lead to different phenomena. For $1 \le d \le n$ let $\eta_{d,n}$ denote the d-th largest of $|X_1|, \ldots, |X_n|$ and let

$$^{(d)}S_n = \sum_{k=1}^{n} X_k I\{|X_k| \le \eta_{d,n}\}. \tag{1.6}$$

If the distribution of X_1 is continuous, then $|X_1|, |X_2|, \ldots$ are different with probability 1, and thus $^{(d)}S_n$ coincides with the usual modulus trimmed sum obtained by discarding from S_n the $d-1$ elements with the largest moduli. Griffin and Pruitt (1987) showed that if X_1 has a symmetric distribution, then $^{(d_n)}S_n$ is asymptotically normal for any $d_n \to \infty$, $d_n/n \to 0$, but this is generally false in the nonsymmetric case. The purpose of this paper is to describe the asymptotic distribution of $^{(d_n)}S_n$ in the general case. Put

$$H(t) = P(|X| \ge t) \quad \text{and} \quad m(t) = \mathsf{E}XI\{|X| \le t\},$$

and let $H^{-1}(t) = \inf\{x : H(x) \le t\}$ $(0 < t < 1)$ denote the generalized inverse of H. Our main result is the following.

Theorem 1.1. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with distribution function F satisfying (1.2), (1.3) and assume that (1.5) holds. Then we have*

$$\frac{1}{A_n} \sum_{i=1}^{[nt]} \big(X_i I\{|X_i| \le \eta_{d,n}\} - m(\eta_{d,n})\big) \xrightarrow{\mathcal{D}[0,1]} W(t) \tag{1.7}$$

where

$$A_n^2 = \frac{\alpha}{2-\alpha} d (H^{-1}(d/n))^2 \tag{1.8}$$

and W is the Wiener process.

Theorem 1.1 shows that allowing a random centering factor, the modulus trimmed CLT holds for continuous i.i.d. variables under exactly the same conditions as under ordinary trimming. If F is not continuous, the sample $(X_1, \ldots, X_n)$ may contain equal elements with positive probability; according to the definition in Griffin and Pruitt (1987), 'ties' between elements with equal moduli are broken according to the order in which the variables occur in $(X_1, \ldots, X_n)$. But no matter how we break the ties, it may happen that from a set of sample elements with equal moduli some are discarded and others are not, which is rather unnatural from the statistical point of view, since trimming is mainly used to improve the performance of statistical procedures by removing large elements from the sample. The definition of $^{(d)}S_n$ in (1.6) resolves this difficulty and leads to satisfactory asymptotic results in the general case.

Theorem 1.1 enables one to give, among others, change point tests for heavy tailed processes, while the standard CUSUM test fails under infinite variances. A fairly precise characterization for the modulus trimmed CLT with nonrandom centering and norming factors was given in Berkes and Horváth (2012).

Under additional technical assumptions on the distribution function of X_1 and on the growth speed of d_n, Theorem 1.1 was proved in Berkes et al. (2011) with a fairly complicated argument. The proof of Theorem 1.1 is much simpler and extends to dependent samples as well, as we will show in a subsequent paper. Let

$$\hat{A}_n^2 = \sum_{i=1}^{n} X_i^2 I\{|X_i| \le \eta_{d,n}\} - \frac{1}{n}\left(\sum_{i=1}^{n} X_i I\{|X_i| \le \eta_{d,n}\}\right)^2.$$

Berkes et al. (2011) showed that under the conditions of Theorem 1.1 we have that

$$\hat{A}_n / A_n \xrightarrow{P} 1$$

and therefore Theorem 1.1 yields

$$\frac{1}{\hat{A}_n}\left(\sum_{i=1}^{[nt]} X_i I\{|X_i| \le \eta_{d,n}\} - \frac{[nt]}{n}\sum_{i=1}^{n} X_i I\{|X_i| \le \eta_{d,n}\}\right) \xrightarrow{\mathcal{D}[0,1]} B(t),$$