Statistics and Probability Letters 135 (2018) 70–75

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Modulo 1 limit theorems for autoregressive random variables and their sums

Bruno Massé

ULCO, LMPA J. Liouville, B.P. 699, F-62228 Calais, France

ARTICLE INFO

Article history:
Received 19 June 2017
Accepted 29 November 2017
Available online 11 December 2017

MSC 2010:
60B10
11K06
60F99
42A16

Keywords:
Modulo 1 distribution
Fractional part
Uniform distribution
Fourier coefficients
Autoregressive process

ABSTRACT

We extend Lévy's and Robbins's results on distribution modulo 1 of sums of i.i.d. random variables to the case of autoregressive random variables and their sums. We show that the results may and may not depend on the regression coefficient and on the distribution of the white noise.

1. Introduction

Given a sequence $Z_n(\omega)$ of random variables (r.v. in abbreviated notation), $\omega \mapsto Z_n(\omega)$ is called below a *time sample* and $n \mapsto Z_n(\omega)$ (ω fixed) is called a *trajectory*.

When $Z_n(\omega)$ is the fractional part of the sum of n independent and identically distributed r.v.'s, Lévy (1939) and Robbins (1953) (see also Berger and Hill (2011, p. 104)) proved that, except for a few special cases, the distribution of the time sample of the process converges weakly to the uniform distribution on $[0, 1)$ and that the trajectories of the process are almost surely uniformly distributed modulo 1 (u.d. mod. 1 in abbreviated form; see the definition below). Some results on the convergence rate for the latter topic are given in Schatte (1988). When $Z_n(\omega)$ is the fractional part of the sum of n independent r.v.'s, Miller and Nigrini (2008) provide in some necessary and sufficient conditions on characteristic functions ensuring that the density of the time sample converges in L_1-norm to the uniform density on $[0, 1)$. See Miller and Nigrini (2008) and Schatte (1988) for more references on the subject.

In the present note, we investigate the distribution modulo 1 of time samples and trajectories of autoregressive processes of order 1 and of partial sums of their terms, thus extending Lévy's and Robbins's results and complementing Miller and Nigrini's ones. Our setting may seem narrow but it suffices to depict the kind of difficulties which arise when considering arrays instead of sequences and when the r.v.'s involved in the sums are not supposed to be identically distributed (see also Example 2.4 in Miller and Nigrini (2008) for the latter topic). We provide a few possible ways of extension in Section 4. Some of our results depend strongly on the value of the regression coefficient and on the distribution of the white noise, others can be described as invariance principles.

E-mail address: bmasse@lmpa.univ-littoral.fr.

https://doi.org/10.1016/j.spl.2017.11.015
0167-7152/

Besides its role in Analytic Number Theory (Chandrasekharan, 1968; Ganville and Rudnick, 2007; Massé and Schneider, 2014), distribution modulo 1 is closely related with *first digit frequency*, *mantissa distribution* and *Benford's law* (Berger and Hill, 2011; Massé and Schneider, 2012), whose scope of application has grown substantially during the past few years (Miller, 2015).

The rest of this section is devoted to the precise description of our framework and to some notation and definitions. Section 2 presents some known tools used in our proofs. Our results are stated and proved in Section 3. In this note, the lemmas are known results and the propositions and the theorems are new.

1.1. The two processes under investigation

We shall consistently use the following notation through this note:

- the sequence $(X_n)_{n\geq 0}$ is defined by $X_n = aX_{n-1} + \varepsilon_n \quad (n \geq 1)$ where $a \neq 1$ is a positive real number, $(\varepsilon_n)_{n\geq 1}$ is a sequence of independent and identically distributed r.v.'s and X_0 is a r.v. independent of $(\varepsilon_n)_{n\geq 1}$;
- we set $Y_0 = X_0$ and $Y_n = X_0 + \cdots + X_n \quad (n \geq 1)$.

Hence

$$X_n = a^n X_0 + a^{n-1}\varepsilon_1 + \cdots + a\varepsilon_{n-1} + \varepsilon_n \tag{1}$$

and

$$Y_n = b_n X_0 + b_{n-1}\varepsilon_1 + \cdots + b_1\varepsilon_{n-1} + \varepsilon_n. \tag{2}$$

where $b_n = \sum_{l=0}^{l=n} a^l$.

We are interested in the asymptotic distribution of the fractional part of X_n and of Y_n and in the distribution of the corresponding trajectories (see the definition in Section 1.2). This framework differs somehow from Miller and Nigrini's one because it does not exist any sequence of independent r.v.'s $(T_n)_{n\geq 0}$ such that $X_n = T_n - T_{n-1}$ or $Y_n = T_n - T_{n-1}$ for all n.

1.2. Notation and definitions

The fractional part $\{x\}$ of a real number x is defined by $\{x\} = x - \lfloor x \rfloor$ where $\lfloor x \rfloor$ is the largest integer less than x. When dealing with fractional parts, we get benefit in distinguishing between the interval $[0, 1)$, endowed with the natural metric, and the *torus* $[0, 1)$ which is endowed with the metric $\Delta(z, t) = \min(|x - y| : \{x\} = z, \{y\} = t)$. This means for example that if $(y_n)_n \subset [0, 1)$ converges to 1 when it is considered as a sequence of real numbers, then it converges to 0 when it is considered as a sequence of fractional parts.

For simplicity, U will designate both the uniform distribution on $[0, 1)$ and any r.v. distributed following it. We use the standard notation: $e_\lambda(x)$ for $\exp(2i\pi\lambda x)$ with $i^2 = -1$ and $\mathbb{Z}^+$ for the set of positive integers. For convenience of exposition, we define the characteristic function of a r.v. Z by $\varphi_Z(t) = \mathbb{E}(e_t(Z))$ (instead of the standard $\varphi_Z(t) = \mathbb{E}(\exp(itZ))$). The common characteristic function of the r.v.'s ε_n and that of X_0 will be denoted φ and φ_0 respectively.

Given a probability measure μ on $[0, 1)$, a sequence $(x_n)_{n\geq 1} \subset [0, 1)$ is said to be *distributed following* μ if, for every μ-continuity point $t \in [0, 1)$,

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{[0,t)}(x_n) = \mu([0, t)). \tag{3}$$

In particular, a sequence $(y_n)_n$ of real numbers is said to be *u.d. mod.* 1 if

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{[0,t)}(\{y_n\}) = t \quad (0 \leq t < 1). \tag{4}$$

Note that (3) is equivalent to the weak convergence to μ of the sequence of probability measures $\left(\frac{1}{N}\sum_{n=1}^{N}\delta_{x_n}\right)_n$ where δ_x denotes the Dirac measure at x.

2. Preliminaries

First, we mention the elementary formula

$$e_h(x) = e_h(\{x\}) \quad (x \text{ real and } h \text{ integer}) \tag{5}$$

which is crucial in all what follows. We collect here the tools used in our proofs. They are known results of *Uniform Distribution Theory* and of *Fourier Analysis*. The next lemma is known as the *Weyl's Criterion* (Kuipers and Niederreiter, 2006, p. 7).