Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications ▮ (▮▮▮▮) ▮▮▮–▮▮▮

stochastic processes and their applications

www.elsevier.com/locate/spa

Products of random variables and the first digit phenomenon

Nicolas Chenavier, Bruno Massé*, Dominique Schneider

Univ. Littoral Côte d'Opale, EA 2597 — Laboratoire de mathématiques pures et appliquées Joseph Liouville, F-62228 Calais, France

Received 28 March 2017; received in revised form 17 July 2017; accepted 7 August 2017
Available online xxxx

Abstract

We provide conditions on dependent and on non-stationary random variables X_n ensuring that the mantissa of the sequence of products $\left(\prod_1^n X_k\right)$ is almost surely distributed following Benford's law or converges in distribution to Benford's law. This is achieved through proving new generalizations of Lévy's and Robbins's results on distribution modulo 1 of sums of independent random variables.

MSC 2010: 60B10; 11B05; 11K99

Keywords: Benford's law; Density; Mantissa; Weak convergence

1. Introduction

Let $b > 1$. *Benford's law in base b* is the probability measure μ_b on the interval $[1, b[$ defined by

$$\mu_b([1, a[) = \log_b a \qquad (1 \le a < b),$$

where $\log_b a$ denotes the logarithm in base b of a. The *mantissa* in base b of a positive real number x is the unique number $\mathcal{M}_b(x)$ in $[1, b[$ such that there exists an integer k satisfying $x = \mathcal{M}_b(x) b^k$.

* Corresponding author.
E-mail addresses: nicolas.chenavier@lmpa.univ-littoral.fr (N. Chenavier), bruno.masse@lmpa.univ-littoral.fr (B. Massé), dominique.schneider@lmpa.univ-littoral.fr (D. Schneider).

http://dx.doi.org/10.1016/j.spa.2017.08.003
0304-4149/

When a sequence of positive random variables (X_n) is of a type usually considered by probabilists and statisticians, there is little to be said on $(\mathcal{M}_b(X_n))$ (see Remark 2.1 in Section 2.2 for instance) while by contrast there is much to report on $\left(\mathcal{M}_b\left(\prod_1^n X_k\right)\right)$ as we will see. Our purpose is therefore to exhibit conditions on X_n ensuring that the sequence $\left(\mathcal{M}_b\left(\prod_1^n X_k\right)\right)$ is almost surely distributed following μ_b (see Definition 2.1) or ensuring that the law of $\mathcal{M}_b\left(\prod_1^n X_k\right)$ converges weakly to μ_b as $n \to +\infty$. We hope that this will enlarge, to a certain extent, the field of applications of the Benford's law (see Section 1.1 for examples of such applications).

To the best of our knowledge, apart from [27], the known results on the asymptotic behaviour of $\left(\mathcal{M}_b\left(\prod_1^n X_k\right)\right)$ only deal with the cases where the X_n are independent and identically distributed and the situations where $X_n = X$ for $n \geq 1$ and X is some random variable (see Section 1.3 for details).

1.1. The first digit phenomenon

Benford [2] noticed in 1938 that many real-life lists of numbers have a strange property: numbers whose mantissae are small are more numerous than those whose mantissae are large. This fact is called the *First Digit Phenomenon*. He also noticed that this phenomenon seems independent of the units. This led him to make a scale-invariance hypothesis (more or less satisfied in real life) from which he derived that μ_{10} can be seen as the (ideal) distribution of digits or mantissa of many real-life numbers. Of course, this ideal distribution is never achieved in practice.

Several mathematicians have been involved in this subject and have provided sequences of positive numbers whose mantissae are (or approach to be) distributed following μ_b in the sense of the natural density [1,7,9,12,25] (see Definition 2.1), random variables whose mantissa law is or approaches μ_b [3,11,16,19,22], sequences of random variables whose mantissae laws converge to μ_b or whose mantissae are almost surely distributed following μ_b [24,27,31,36]. Some convergence rates for the mantissa of products of i.i.d. random variables are provided by Schatte in [33,35]. The same author gives in [34] a survey on mantissa distribution and conjectures that extensive computing leads to numbers whose mantissae are close to be distributed following μ_B.

Among the many applications of the First Digit Phenomenon, we can quote: fraud detection [29], computer design [15,20] (data storage and roundoff errors), image processing [37] and data analysis in natural sciences [28,32]. See [5,26] for more details.

1.2. Content

Section 2 is devoted to notation, definitions and tools from Uniform Distribution Theory. Our main results, Theorems 3.3 and 3.9, are presented in Section 3. They both involve $T_N^{(h)} := (1/N)\sum_{n=1}^{N} \exp\left(2i\pi h \left(\sum_1^n \log_b X_k\right)\right)$. Theorem 3.3 states that, under the assumption that the sequence (X_n) is stationary, the sequence $\left(\mathcal{M}_b\left(\prod_1^n X_k\right)\right)$ is almost surely distributed following μ_b if and only if, for every positive integer h, $\mathbb{E}T_N^{(h)}$ converges to 0 as $N \to +\infty$. Theorem 3.9 states that the following condition, without constraints on the dependence and on the stationarity of the X_n, is sufficient: for every positive integer h, $\sum_{N=1}^{\infty} \frac{\mathbb{E}|T_N^{(h)}|^p}{N} < +\infty$ for some $p \geq 1$. These properties are used in Section 4 to investigate the cases where the random variables X_n are stationary and log-normal, are exchangeable, are stationary and 1-dependent and the case where they are independent and non-stationary. We provide in the Appendix a survey of the main known properties of Benford's law (scale-invariance, power-invariance and invariance under mixtures). We think that this might help, together with Section 1.3, to put our results in perspective.

Please cite this article in press as: N. Chenavier, et al., Products of random variables and the first digit phenomenon, Stochastic Processes and their Applications (2017), http://dx.doi.org/10.1016/j.spa.2017.08.003.