Statistics and Probability Letters 97 (2015) 229–240

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Descents following maximal values in samples of geometric random variables

Margaret Archibald [1], Aubrey Blecher, Charlotte Brennan [*,2], Arnold Knopfmacher [3]

The John Knopfmacher Centre for Applicable Analysis and Number Theory, School of Mathematics, University of the Witwatersrand, Private Bag 3, Wits 2050, Johannesburg, South Africa

ARTICLE INFO

Article history:
Received 8 August 2014
Received in revised form 13 October 2014
Accepted 28 November 2014
Available online 7 December 2014

MSC:
primary 05A15
05A16
secondary 60C05

Keywords:
Geometric random variable
Generating function
Rice's method
Asymptotic approximation
Descent
Maximum

ABSTRACT

We consider samples of geometric random variables and find the average size of the descent after the first and last maximal values. These are asymptotically but not exactly equal, with the descent after the last maximum being slightly larger than that after the first. Thereafter we calculate the probability that the descent after the last maximum is equal to, greater than, or less than the descent after the first maximum. Finally we compute asymptotic expansions for these probabilities.

1. Introduction

A sample of geometric random variables of length n is a list of independent and identically distributed geometric random variables $(\Gamma_1, \Gamma_2, \ldots, \Gamma_n)$ where $\mathbb{P}\{\Gamma_j = i\} = pq^{i-1}$, for $1 \le j \le n$, with $p + q = 1$. The smaller the value of q, the greater the prevalence of smaller numbers in the sample.

Samples of geometric random variables have been conceptualised in many different ways; we provide two here. The first is in terms of a coin-flipping game where each coin represents an independent variable. We set q to be the probability of getting "heads", and p to be the probability of getting "tails". The idea is that n coins labelled $1, \ldots, n$ are flipped sequentially and if a coin shows a head then it will be flipped again in the next round, but if it shows a tail it will not be flipped again and the value of the kth random variable is the number of times the kth coin is flipped. A maximum value will be the number of times we flip a coin that makes it to the last round.

* Corresponding author.
E-mail addresses: Margaret.Archibald@wits.ac.za (M. Archibald), Aubrey.Blecher@wits.ac.za (A. Blecher), Charlotte.Brennan@wits.ac.za (C. Brennan), Arnold.Knopfmacher@wits.ac.za (A. Knopfmacher).

1 This material is based upon work supported by the National Research Foundation under grant number 89147.
2 This material is based upon work supported by the National Research Foundation under grant number 86329.
3 This material is based upon work supported by the National Research Foundation under grant number 81021.

http://dx.doi.org/10.1016/j.spl.2014.11.023

Alternatively, in the urn model we consider an infinite sequence of urns (or boxes) one behind the other. Stand at one end and throw n balls labelled $1, \ldots, n$ towards the urns. If we label the urns $1, 2, 3, \ldots$ then the probability that ball labelled k falls into the ith urn is pq^{i-1}. In this case the kth geometric random variable is set as i. A maximum value in this case will be obtained from any ball that lies in the furthest urn after all the balls have been thrown.

A maximum value (there may be more than one) in a sample of geometric random variables has a value which is larger or equal to any other value in the sample. We refer to the left-most of these largest values as the first maximum, and the right-most as the last maximum.

Some relevant references on samples of geometric random variables and their applications can be found in Archibald et al. (2006), Grabner et al. (2000), Kirschenhofer and Prodinger (1990), Kirschenhofer and Prodinger (1993), Louchard and Prodinger (2008), Prodinger (1996) and Szpankowski and Rego (1990). For papers relating to maxima particularly, see Archibald (2005), Archibald and Knopfmacher (2007), Archibald and Knopfmacher (2009), Baryshnikov et al. (1995), Eisenberg et al. (1993), Kirschenhofer and Prodinger (1996), Knopfmacher and Prodinger (2004), Prodinger (1996), and Brennan and Knopfmacher (2005a), Brennan and Knopfmacher (2005b), Knopfmacher and Prodinger (2004), Knopfmacher and Prodinger (2006), Knopfmacher and Prodinger (2007), Louchard and Prodinger (2005) relate to descents and ascents in these samples.

In this paper we are interested in the size of the descent after the first (resp. last) maximum in a sample of geometric random variables. By "descent" we mean the difference in size between a value and its right-hand neighbour.

The related question of separation of the maxima was recently studied in Brennan et al. (2011). The results there show that maxima tend to be far apart, suggesting that a large descent should follow a maximum value. We quantify this in the present paper.

If there is a maximum (with value h) in the last position, we say it is followed by a descent of size h.

Example. If $\Gamma = \Gamma_1, \ldots, \Gamma_n = 211412431$ is a sample of geometric random variables of size (length) $n = 9$, then the first maximum occurs in position 4 and the last maximum occurs in position 7. The descent after the first maximum is 3 and the descent after the last maximum is 1.

Analogous problems have recently been considered in the case of compositions of the integer n, see Blecher et al. (2014).

2. Notation

The symbolic notation used in this paper is described here for easy reference.

$$Q := 1/q$$
$$p = 1 - q$$
$$L := \log Q \ (= \ln Q)$$
$$\chi_k := \frac{2k\pi \mathrm{i}}{L}, \quad \text{where } \mathrm{i} = \sqrt{-1};\ k \in \mathbb{Z} \text{ and } k \neq 0$$
$$H_k := \sum_{i=1}^{k} \frac{1}{i} \quad \text{(denotes the kth harmonic number)}$$
$$\gamma := 0.5772\ldots \quad \text{(denotes Euler's constant)}$$
$$\Gamma(z) := \int_0^\infty t^{z-1} e^{-t} dt \quad \text{for Re } (z) > 0$$
$$\Gamma(n) := (n-1)! \quad \text{for } n \in \mathbb{Z},\ n \geq 1$$
$$[n-1; z] := \frac{\Gamma(n)\Gamma(-z)}{\Gamma(n-z)}.$$

In the following, Ø stands for "possibly empty"; B stands for "block", and E stands for "element" (a single geometric random variable).

i. E_h := an element of size h.
ii. $E_{<h}$:= an element of size $h - 1$ or smaller.
iii. $B^{\emptyset}_{<h}$:= a block of letters of size $h - 1$ or less which may possibly be empty.
iv. $B_{<h}$:= a block of letters of size $h - 1$ or less which has at least one element in it.

3. Descent after the first maximum

In this section, we calculate the average size of the descent after the first maximum in a sample of geometric random variables. The results are as follows: