МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2021
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2021
Proceedings of the Fiftieth Spring Conference
of the Union of Bulgarian Mathematicians
2021

A LIMIT THEOREM FOR THE MAXIMUM OF RANDOM VARIABLES WITH LOGARITHMIC DISTRIBUTION*

Kosto V. Mitov

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables and let $M_n = \max(X_1, \ldots, X_n)$ denote their maximum. In the case of discrete distribution of X_i an appropriate normalization for a non-degenerate limiting distribution of M_n does not always exists. In such a case is usually used the scheme of series. In this note, we derive a non-degenerate limiting distribution for the maximum of random variables with logarithmic distribution in the scheme of series.

1. Introduction. Let X be a discrete random variable taking the non-negative integers with probability mass function (pmf) $\Pr(X = k) = p_k$ and cumulative distribution function (cdf) F. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed (iid) copies of X and let us denote the maximum by $M_n = \max(X_1, \ldots, X_n)$. Consider the limiting distribution of $(M_n - b(n))/a(n)$ as $n \to \infty$ for some real numbers $a(n) > 0$ and $b(n)$. It is well known that there are no $a(n) > 0$ and $b(n)$ providing a proper limiting distribution if

$$\frac{\Pr(X = k)}{1 - F(k-1)} = p_k \Big/ \sum_{j=k}^{\infty} p_j \tag{1}$$

fails to converge to 0 as $k \to \infty$ (see (Corollary 2.4.1, Galambos, 1987).

It is known that for distributions such as Poisson and geometric distribution the condition (1) is not valid. So for these and some other discrete distributions non-degenerate limiting distributions do not exist. Considering the maxima for such random variables in a scheme of series where the parameters of the distribution vary it is possible to obtain proper limits under linear normalization (see Anderson [1], Anderson et al.[2] for Poisson distribution, and Nadarajah and Mitov [4, 5] for geometric, binomial, negative binomial). All these distributions belong to the class of power series distributions.

The logarithmic distribution also belongs to this class. This distribution is quite popular in insurance for modeling a claim frequency. It has been used to describe, for example: the number of items purchased by a consumer in a particular period; the number of bird and plant species in an area; and the number of parasites per host.

---

*The research was partially supported by the NFSR of the Min. Edu. Sci. of Bulgaria, grant No. KP-6-N22/3.

2020 Mathematics Subject Classification: 60F05.

Key words: Extremes, Limit theorem, Linear normalization, Logarithmic distribution.

So, it is interesting from a pure and an applied point of view to obtain non degenerate limiting distributions for the maxima of logarithmic distribution. Unfortunately, similarly to the other power series distributions, it does not satisfy the condition (1).

In this note we use the scheme of series varying the parameter of the distribution in order to obtain a proper limiting distribution.

2. The results. The next lemma is used in the proof of the main result.

Lemma 1. *For two positive functions such that* $a(x) \to \infty$, $b(x) \to 0$, *and* $a(x)\, b(x) \to 0$ *as* $x \to \infty$,

$$(1 + b(x))^{\frac{a(x)}{b(x)}} \sim \exp\left(a(x)\right), \ x \to \infty.$$

Proof. Indeed

$$\frac{(1 + b(x))^{\frac{a(x)}{b(x)}}}{e^{a(x)}} = \exp\left(\frac{a(x)}{b(x)} \log\left(1 + b(x)\right) - a(x)\right).$$

Using Taylor expansion for $\log(1 + x) = x - \frac{x^2}{2}(1 + o(1))$ as $x \to 0$, we have

$$\begin{aligned} &\frac{a(x)}{b(x)} \log\left(1 + b(x)\right) - a(x) \\ = \ & a(x)\left[\frac{1}{b(x)}\left(b(x) - \frac{b(x)^2}{2}(1 + o(1))\right) - 1\right] \\ = \ & -\frac{a(x)b(x)}{2}(1 + o(1)) \to 0, \ x \to \infty. \end{aligned}$$

The proof is complete. □

We will also use the following result (Theorem 1.5.1, Leadbetter et al., [3]). If $a(n) > 0$ and $b(n)$ are sequences of real numbers such that

(2) $$n\left\{1 - F\left(a(n)x + b(n)\right)\right\} \to -\log H(x), \qquad n \to \infty$$

then

(3) $$\Pr\left\{M_n \le a(n)x + b(n)\right\} \to H(x), \qquad n \to \infty.$$

The logarithmic distribution has a probability mass function (pmf) and a cumulative distribution function (cdf) as follows

$$p_k = -\frac{1}{\log(1 - p)} \frac{p^k}{k}, \ p \in (0, 1), \ k = 1, 2, \ldots$$

and

$$F(x) = \Pr(X \le x) = -\frac{1}{\log(1 - p)} \sum_{j=1}^{[x]} \frac{p^j}{j}, \quad x \ge 0,$$

respectively. We can write

$$\begin{aligned} 1 - F(k - 1) &= 1 + \frac{1}{\log(1 - p)} \sum_{j=1}^{k-1} \frac{p^j}{j} = -\frac{1}{\log(1 - p)} \sum_{j=k}^{\infty} \frac{p^j}{j} \\ &= -\frac{1}{\log(1 - p)} p^k \sum_{j=0}^{\infty} \frac{p^j}{j + k}. \end{aligned}$$

Now we can see that

$$\frac{p_k}{1-F(k-1)} = \frac{-\dfrac{1}{\log(1-p)}\dfrac{p^k}{k}}{-\dfrac{1}{\log(1-p)}p^k\displaystyle\sum_{j=0}^{\infty}\frac{p^j}{j+k}} = \frac{1}{\displaystyle\sum_{j=0}^{\infty}\frac{p^j}{1+\frac{j}{k}}}$$

$$= \frac{1}{\displaystyle\sum_{j=0}^{\infty}p^j - \sum_{j=0}^{\infty}p^j\frac{\frac{j}{k}}{1+\frac{j}{k}}} = \frac{1}{S_1-S_2},$$

where $S_1 = \dfrac{1}{1-p}$ and

$$0 \le S_2 \le \frac{1}{k}\sum_{j=1}^{\infty} jp^j = \frac{p}{k}\frac{1}{(1-p)^2} \to 0, \quad k\to\infty.$$

Therefore,

$$\frac{p_k}{1-F(k-1)} \to 1-p \in (0,1).$$

The condition (1) does not hold. Hence, there are no normalizing sequences $a(n)$ and $b(n)$ such that $(M_n - b(n))/a(n)$ converges to a non-degenerate max stable distribution. Now we let $p = p(n) \to 1$, as $n\to\infty$ in order to find appropriate normalizing sequences. The following theorem states the result.

Theorem 1. *Let $X_1, X_2, \ldots, X_n$ be iid logarithmic random variables a with parameter $p = p(n)$. If $p(n)\to 1$ according to*

$$q(n) := 1-p(n) = o\left(\frac{1}{\log n}\right) \text{ and } \log(1/q(n)) = o(n), \quad n\to\infty,$$

there are sequences

$$a(n) = \frac{\alpha}{q(n)}, \text{ where } \alpha>0 \text{ is fixed, and } \quad b(n) = \frac{\log n - \log\log n - \log\log\frac{1}{q(n)}}{q(n)},$$

such that for $x>0$,

$$\lim_{n\to\infty}\Pr\left\{\frac{M_n-b(n)}{a(n)} \le x\right\} = \exp\{-\exp(-\alpha x)\}.$$

Remark 1. *The conditions for $q(n)$ are satisfied if for example $q(n)$ varies regularly, i.e. $q(x) = x^\theta L(x)$, $x\to\infty$, $\theta\le 0$. The second condition yields the relation $\log\log\dfrac{1}{q(n)} = o(\log n)$, $n\to\infty$, which is used in the proof.*

Proof. We have

$$\begin{aligned}
& n\{1-F_n(a(n)x+b(n))\} \\
= & -\frac{n}{\log q(n)}\sum_{j=[a(n)x+b(n)]+1}^{\infty}\frac{(1-q(n))^j}{j} \\
= & -\frac{n}{\log q(n)}(1-q(n))^{[a(n)x+b(n)]}\sum_{j=1}^{\infty}\frac{(1-q(n))^j}{j+[a(n)x+b(n)]}.
\end{aligned}$$

Using the inequalities $x - 1 \leq [x] \leq x$, we have

$$n\left\{1 - F_n\left(a(n)x + b(n)\right)\right\} \leq -\frac{n}{\log q(n)}\left(1 - q(n)\right)^{a(n)x+b(n)-1}\sum_{j=1}^{\infty}\frac{(1-q(n))^j}{j + a(n)x + b(n) - 1}$$

$$= -\frac{n}{\log q(n)}\frac{\left(1 - q(n)\right)^{a(n)x+b(n)-1}}{b(n)}\sum_{j=1}^{\infty}\frac{(1-q(n))^j}{1 + \frac{j+a(n)x-1}{b(n)}} \tag{4}$$

and

$$n\left\{1 - F_n\left(a(n)x + b(n)\right)\right\} \geq -\frac{n}{\log q(n)}\left(1 - q(n)\right)^{a(n)x+b(n)}\sum_{j=1}^{\infty}\frac{(1-q(n))^j}{j + a(n)x + b(n)}$$

$$= -\frac{n}{\log q(n)}\frac{\left(1 - q(n)\right)^{a(n)x+b(n)}}{b(n)}\sum_{j=1}^{\infty}\frac{(1-q(n))^j}{1 + \frac{j+a(n)x}{b(n)}}. \tag{5}$$

Let us consider the right-hand side of (5) first. We have

$$\Pi_n(x) = -\frac{n}{\log q(n)}\frac{\left(1 - q(n)\right)^{a(n)x+b(n)}}{b(n)} = -\left(1 - q(n)\right)^{a(n)x}\frac{n}{\log q(n)}\frac{\left(1 - q(n)\right)^{b(n)}}{b(n)},$$

$$\left(1 - q(n)\right)^{a(n)x} = \left(1 - q(n)\right)^{\alpha x/q(n)} \to \exp(-\alpha x), \quad n \to \infty$$

and

$$\frac{n}{\log q(n)}\frac{\left(1 - q(n)\right)^{b(n)}}{b(n)} = \frac{n}{\log q(n)}\frac{q(n)}{\log n - \log\log n - \log\log\left(1/q(n)\right)}$$
$$\cdot \left(1 - q(n)\right)^{\frac{\log n}{q(n)}}\left(1 - q(n)\right)^{\frac{-\log\log n}{q(n)}}\left(1 - q(n)\right)^{-\frac{\log\log(1/q(n))}{q(n)}}.$$

Therefore, using $q(n)\log n \to 0$, $n \to \infty$ we obtain by Lemma 1:

$$\left(1 - q(n)\right)^{\frac{\log n}{q(n)}} \sim e^{-\log n} \sim \frac{1}{n}, \quad n \to \infty,$$
$$\left(1 - q(n)\right)^{\frac{-\log\log n}{q(n)}} \sim e^{\log\log n} \sim \log n, \quad n \to \infty,$$
$$\left(1 - q(n)\right)^{-\frac{\log\log(1/q(n))}{q(n)}} \sim e^{\log\log(1/q(n))} \sim \log\left(1/q(n)\right) \sim -\log q(n), \quad n \to \infty.$$

These relations provide

$$\Pi_n(x) \sim q(n)e^{-\alpha x}, \quad n \to \infty. \tag{6}$$

Now we will prove that

$$S(n) = q(n)\sum_{j=1}^{\infty}\frac{(1-q(n))^j}{1 + \frac{j+a(n)x}{b(n)}} \to 1, \quad n \to \infty.$$

We have the following estimate:

$$\begin{aligned} S(n) &= q(n)\sum_{j=1}^{\infty}(1-q(n))^j\left(1 - \frac{\frac{j+a(n)x}{b(n)}}{1 + \frac{j+a(n)x}{b(n)}}\right) \\ &= q(n)\sum_{j=1}^{\infty}(1-q(n))^j - q(n)\sum_{j=1}^{\infty}(1-q(n))^j\frac{\frac{j+a(n)x}{b(n)}}{1 + \frac{j+a(n)x}{b(n)}} \\ &= S_1(n) - S_2(n). \end{aligned}$$

For $S_1(n)$, we obtain ($0 < q(n) < 1$ for any fixed n)

$$S_1(n) = q(n)\frac{1-q(n)}{1-(1-q(n))} = 1 - q(n) \to 1, \quad n \to \infty.$$

For $S_2(n)$ we obtain

$$\begin{aligned}
0 &\leq S_2(n) \\
&= q(n)\sum_{j=1}^{\infty}(1-q(n))^j \frac{\frac{j+a(n)x}{b(n)}}{1+\frac{j+a(n)x}{b(n)}} \\
&\leq q(n)\sum_{j=1}^{\infty}(1-q(n))^j \frac{j+a(n)x}{b(n)} \\
&= \frac{q(n)}{b(n)}\sum_{j=1}^{\infty} j\,(1-q(n))^j + \frac{q(n)a(n)x}{b(n)}\sum_{j=1}^{\infty}(1-q(n))^j \\
&= \frac{q(n)}{b(n)}\frac{1-q(n)}{(1-(1-q(n)))^2} + \frac{q(n)a(n)x}{b(n)}\frac{1-q(n)}{1-(1-q(n))} \\
&= \frac{1-q(n)}{b(n)q(n)} + \frac{(1-q(n))\,a(n)x}{b(n)} \to 0, \quad n \to \infty
\end{aligned}$$

because under the conditions of the theorem $q(n)b(n) = \log n\,(1+o(1))$ and $a(n)/b(n) \to 0$ as $n \to \infty$. Therefore

$$S(n) \to 1, \quad n \to \infty.$$

The last limit and (6) provide

$$\liminf_{n\to\infty} n\left\{1 - F_n\left(a(n)x + b(n)\right)\right\} \geq e^{-\alpha x}. \tag{7}$$

The right-hand side of (4) can be manipulated in the same way to yield

$$\limsup_{n\to\infty} n\left\{1 - F_n\left(a(n)x + b(n)\right)\right\} \leq e^{-\alpha x}. \tag{8}$$

From (7), (8), (4), (5), (2), and (3) we obtain the proof. □

REFERENCES

[1] C. W. Anderson. Extreme value theory for a class of discrete distributions with applications to some stochastic processes. *Journal of Applied Probability*, **7**, (1970) 99–113.

[2] C. W. Anderson, S. G. Coles, J. Hüsler. Maxima of Poisson-like variables and related triangular arrays. *Annals of Applied Probability*, **7** (1997), 953–971.

[3] M. R. Leadbetter, G. Lindgren, H. Rootzén. Extremes and Related Properties of Random Sequences and Processes, New York, Springer Verlag, 1987.

[4] S. Nadarajah, K. V. Mitov. Asymptotics of maxima of discrete random variables. *Extremes*, **5** (2002), 287–294.

[5] S. Nadarajah, K. V. Mitov. Extremal limit laws for discrete random variables. *Journal of Mathematical Sciences*, **122** (2004), 3404–3415.

Kosto V. Mitov
Faculty of Pharmacy
Medical University of Pleven
1, Saint Kliment Ohridski Str.
5800 Pleven, Bulgaria
e-mail: kmitov@yahoo.com

ГРАНИЧНА ТЕОРЕМА ЗА МАКСИМУМ НА СЛУЧАЙНИ ВЕЛИЧИНИ С ЛОГАРИТМИЧНО РАЗПРЕДЕЛЕНИЕ

Косто В. Митов

Нека $X_1, X_2, \ldots, X_n$ са независими, еднакво разпределени случайни величини и $M_n = \max(X_1, \ldots, X_n)$ е техният максимум. Когато разпределението на X_i е дискретно, не винаги съществува линейна нормализация, която осигурява неизродено гранично разпределение на максимума M_n. В такива случаи се използва схема от серии. В тази бележка е намерено неизродено гранично разпределение на M_n, когато случайните величини X_i имат логаритмично разпределение, като е използвана схема от серии.