

STRONG LAWS FOR THE LARGEST RATIO OF ADJACENT ORDER STATISTICS

ANDRÉ ADLER

Department of Mathematics, Illinois Institute of Technology, Chicago, Illinois, 60616, USA.
E-mail: adler@iit.edu

Abstract

Consider independent and identically distributed random variables $\{X_{n,k}, 1 \leq k \leq m_n, n \geq 1\}$. We order this data set, $X_{n(1)} < X_{n(2)} < X_{n(3)} < \cdots < X_{n(m_n-1)} < X_{n(m_n)}$. Then we find the ratio of these adjacent order statistics. Our random variable of interest is the largest of these adjacent ratios, $\max_{2 \leq k \leq m_n} X_{n(k)}/X_{n(k-1)}$. We obtain various limit theorems for this random variable.

1. Introduction

In this paper we establish limit theorems for the largest ratio of adjacent order statistics from a sample of size m_n , where the only restriction on m_n is that we prevent it from going to infinity. We examine the uniform and the Pareto distributions. But in the last section we show how this can be extended to the exponential and many other distributions. In order to do this we first take the joint density of the random sample and then transform that to the order statistics. Then we obtain the joint distribution of the smallest order statistic with the $m_n - 1$ adjacent ratios. Then by integrating out the smallest order statistic we have the joint density of all the adjacent ratios. From that joint density we obtain the distribution of the largest one.

This is similar to what was done in various papers. Initially in [1], sums of order statistics from Pareto random variables were observed. Then in [2], ratios of order statistics from Paretos were examined. Likewise the ratios

Received June 18, 2017 and in revised form November 24, 2017.

AMS Subject Classification: 60F15.

Key words and phrases: Almost sure convergence, strong law of large numbers, order statistics, exact laws.

of these order statistics from the uniform distribution was explored in [3], then later in [7]. The exponential distribution was first covered in [4], then in [6] and also in [8]. What is truly fascinating about these theorems is that the ratio of any of the order statistics to the minimal order statistic was almost the same, which defies logic. And these statistics, especially in the exponential setting is quite important. The exponential distribution measure lifetimes of equipment, so these ratios are informing us as to the stability of these systems.

In these previous cases we only looked at the marginal distributions of these ratios. To obtain the distribution of the largest of these ratios we must examine their joint density. What is remarkable is how we can obtain unusual strong laws for the largest of these ratios, which are known as Exact Strong Laws and also some classic strong laws. In all of our cases the size of our samples plays a minor role in the final result and in the uniform case the value of our parameter, θ_n , is completely unimportant.

As usual, we define $\lg x = \log(\max\{e, x\})$ and $\lg_2 x = \lg(\lg x)$. We use throughout the paper the constant C as a generic real number that is not necessarily the same in each appearance, it is just an upper bound in our calculations. And $I(\cdot)$ is our indicator function.

2. Uniform Distribution

We start with independent and identically distributed uniform random variables where the parameter can change from sample to sample. Let $X_{n,1}, \dots, X_{n,m_n}$ be i.i.d. $U(0, \theta_n)$ random variables. The joint density of the original data is

$$f(x_{n,1}, \dots, x_{n,m_n}) = \frac{1}{\theta_n^{m_n}} I(0 < x_{n,1} < \theta_n) \cdots I(0 < x_{n,m_n} < \theta_n).$$

The joint density of the corresponding order statistics, $X_{n(1)}, \dots, X_{n(m_n)}$, is

$$f(x_{n(1)}, \dots, x_{n(m_n)}) = \frac{m_n!}{\theta_n^{m_n}} I(0 < x_{n(1)} < x_{n(2)} < x_{n(3)} < \cdots < x_{n(m_n)} < \theta_n).$$

Next we obtain the joint density of $X_{n(1)}, R_{n,2}, \dots, R_{n,m_n}$, where $R_{n,k} = X_{n(k)}/X_{n(k-1)}$. In order to do that we need the inverse transformations,

which are

$$X_{n(1)} = X_{n(1)}$$

$$X_{n(2)} = X_{n(1)}R_{n,2}$$

$$X_{n(3)} = X_{n(1)}R_{n,2}R_{n,3}$$

$$X_{n(4)} = X_{n(1)}R_{n,2}R_{n,3}R_{n,4}$$

through

$$X_{n(m_n)} = X_{n(1)}R_{n,2}R_{n,3}R_{n,4} \cdots R_{n,m_n}.$$

In order to obtain this density we need the Jacobian, which is the determinant of the matrix

$$\begin{pmatrix} \frac{\partial x_{n(1)}}{\partial x_{n(1)}} & \frac{\partial x_{n(1)}}{\partial r_{n,2}} & \frac{\partial x_{n(1)}}{\partial r_{n,3}} & \cdots & \frac{\partial x_{n(1)}}{\partial r_{n,m_n}} \\ \frac{\partial x_{n(2)}}{\partial x_{n(1)}} & \frac{\partial x_{n(2)}}{\partial r_{n,2}} & \frac{\partial x_{n(2)}}{\partial r_{n,3}} & \cdots & \frac{\partial x_{n(2)}}{\partial r_{n,m_n}} \\ \frac{\partial x_{n(3)}}{\partial x_{n(1)}} & \frac{\partial x_{n(3)}}{\partial r_{n,2}} & \frac{\partial x_{n(3)}}{\partial r_{n,3}} & \cdots & \frac{\partial x_{n(3)}}{\partial r_{n,m_n}} \\ \frac{\partial x_{n(m_n)}}{\partial x_{n(1)}} & \frac{\partial x_{n(m_n)}}{\partial r_{n,2}} & \frac{\partial x_{n(m_n)}}{\partial r_{n,3}} & \cdots & \frac{\partial x_{n(m_n)}}{\partial r_{n,m_n}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

which is the lower triangular matrix

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ r_{n,2} & x_{n(1)} & 0 & \cdots & 0 \\ r_{n,2}r_{n,3} & x_{n(1)}r_{n,3} & x_{n(1)}r_{n,2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{n,2} \cdots r_{n,m_n} & x_{n(1)}r_{n,3} \cdots r_{n,m_n} & x_{n(1)}r_{n,2}r_{n,4} \cdots r_{n,m_n} & \cdots & x_{n(1)}r_{n,2} \cdots r_{n,m_n-1} \end{pmatrix}.$$

Thus the Jacobian is the product of the diagonal elements, which is

$x_{n(1)}^{m_n-1} r_{n,2}^{m_n-2} r_{n,3}^{m_n-3} r_{n,4}^{m_n-4} \cdots r_{n,m_n-1}$. Hence the joint density of $X_{n(1)}, R_{n,2}, \dots, R_{n,m_n}$ is

$$\begin{aligned} f(x_{n(1)}, r_{n,2}, \dots, r_{n,m_n}) &= \frac{m_n!}{\theta_n^{m_n}} x_{n(1)}^{m_n-1} r_{n,2}^{m_n-2} r_{n,3}^{m_n-3} r_{n,4}^{m_n-4} \cdots r_{n,m_n-1} \\ &\quad \cdot I(r_{n,2} > 1) I(r_{n,3} > 1) I(r_{n,4} > 1) \cdots I(r_{n,m_n} > 1) \\ &\quad \cdot I\left(0 < x_{n(1)} < \frac{\theta_n}{r_{n,2}r_{n,3}r_{n,4} \cdots r_{n,m_n}}\right). \end{aligned}$$

Integrating out the unnecessary random variable, $X_{n(1)}$, we finally have the joint density of our ratios of adjacent order statistics

$$\begin{aligned} & f(r_{n,2}, \dots, r_{n,m_n}) \\ &= \frac{m_n!}{\theta_n^{m_n}} \int_0^{\frac{\theta_n}{r_{n,2}r_{n,3}r_{n,4}\cdots r_{n,m_n}}} x_{n(1)}^{m_n-1} r_{n,2}^{m_n-2} r_{n,3}^{m_n-3} r_{n,4}^{m_n-4} \cdots r_{n,m_n-1} dx_{n(1)} \\ &= \frac{(m_n - 1)! r_{n,2}^{m_n-2} r_{n,3}^{m_n-3} r_{n,4}^{m_n-4} \cdots r_{n,m_n-1}}{\theta_n^{m_n}} \left(\frac{\theta_n}{r_{n,2}r_{n,3}r_{n,4} \cdots r_{n,m_n}} \right)^{m_n} \\ &= (m_n - 1)! r_{n,2}^{-2} r_{n,3}^{-3} r_{n,4}^{-4} \cdots r_{n,m_n-1}^{-m_n} r_{n,m_n}^{-m_n} \\ &\quad \cdot I(r_{n,2} > 1) I(r_{n,3} > 1) I(r_{n,4} > 1) \cdots I(r_{n,m_n} > 1). \end{aligned}$$

This shows that the random variables $R_{n,2}, \dots, R_{n,m_n}$ are independent, but not identically distributed. Let $M_n = \max_{2 \leq k \leq m_n} R_{n,k}$. We can now obtain our desired distribution

$$\begin{aligned} F_{M_n}(a) &= P\{M_n \leq a\} \\ &= P\{R_{n,2} \leq a\} P\{R_{n,3} \leq a\} P\{R_{n,4} \leq a\} \cdots P\{R_{n,m_n} \leq a\} \\ &= (m_n - 1)! \int_1^a r_{n,2}^{-2} dr_{n,2} \int_1^a r_{n,3}^{-3} dr_{n,3} \int_1^a r_{n,4}^{-4} dr_{n,4} \cdots \int_1^a r_{n,m_n}^{-m_n} dr_{n,m_n} \\ &= \left(1 - \frac{1}{a}\right) \left(1 - \frac{1}{a^2}\right) \left(1 - \frac{1}{a^3}\right) \cdots \left(1 - \frac{1}{a^{m_n-1}}\right). \end{aligned}$$

Notice that the parameter, θ_n , does not appear in this distribution. Likewise we need not take the same size sample each time. As long as the sample size does not go to infinity we have $f_{M_n}(a) = a^{-2} + O(a^{-3})$ and from that we can obtain an Exact Strong Law.

Theorem 1. *If $X_{n,1}, \dots, X_{n,m_n}$ are i.i.d. $U(0, \theta_n)$ random variables with $\sup_{n \geq 1} m_n < \infty$, then for all $\alpha > 0$ we have*

$$\lim_{N \rightarrow \infty} \frac{\sum_{n=1}^N \frac{(\lg n)^{\alpha-2}}{n} M_n}{(\lg N)^\alpha} = \frac{1}{\alpha} \quad \text{almost surely.}$$

Proof. Let $a_n = (\lg n)^{\alpha-2}/n$, $b_n = (\lg n)^\alpha$ and $c_n = b_n/a_n = n(\lg n)^2$. We use the usual partition

$$\frac{1}{b_N} \sum_{n=1}^N a_n M_n = \frac{1}{b_N} \sum_{n=1}^N a_n [M_n I(1 \leq M_n \leq c_n) - EM_n I(1 \leq M_n \leq c_n)]$$

$$+ \frac{1}{b_N} \sum_{n=1}^N a_n M_n I(M_n > c_n) + \frac{1}{b_N} \sum_{n=1}^N a_n EM_n I(1 \leq M_n \leq c_n).$$

The first term vanishes almost surely by the Khintchine-Kolmogorov Convergence Theorem, see page 113 of [5], and Kronecker’s lemma since

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{c_n^2} EM_n^2 I(1 \leq M_n \leq c_n) &\leq C \sum_{n=1}^{\infty} \frac{1}{c_n^2} \int_1^{c_n} dx \leq C \sum_{n=1}^{\infty} \frac{1}{c_n} \\ &= C \sum_{n=1}^{\infty} \frac{1}{n(\lg n)^2} < \infty. \end{aligned}$$

The second term vanishes, with probability one, by the Borel-Cantelli lemma since

$$\sum_{n=1}^{\infty} P\{M_n > c_n\} \leq C \sum_{n=1}^{\infty} \frac{1}{c_n} < \infty.$$

As for the third term in our partition, we have

$$EM_n I(1 \leq M_n \leq c_n) \sim \int_1^{c_n} \frac{dx}{x} = \lg c_n \sim \lg n.$$

Thus

$$\frac{\sum_{n=1}^N a_n EM_n I(1 \leq M_n \leq c_n)}{b_N} \sim \frac{\sum_{n=1}^N \frac{(\lg n)^{\alpha-1}}{n}}{(\lg N)^\alpha} \rightarrow \frac{1}{\alpha}$$

which completes the proof. □

3. Pareto Distribution

In our next case we observe m_n i.i.d. random variables where the underlying density is $f(x) = px^{-p-1}I(x \geq 1)$, with $p > 0$. The joint density of the original data is

$$f(x_{n,1}, \dots, x_{n,m_n}) = p^{m_n} x_{n,1}^{-p-1} \dots x_{n,m_n}^{-p-1} I(x_{n,1} > 1) \dots I(x_{n,m_n} > 1).$$

The joint density of the corresponding order statistics, $X_{n(1)}, \dots, X_{n(m_n)}$, is

$$\begin{aligned} &f(x_{n(1)}, \dots, x_{n(m_n)}) \\ &= m_n! p^{m_n} x_{n(1)}^{-p-1} \dots x_{n(m_n)}^{-p-1} I(1 < x_{n(1)} < x_{n(2)} < \dots < x_{n(m_n)}). \end{aligned}$$

The Jacobian is the same as in the previous section, hence the joint density of $X_{n(1)}, R_{n,2}, \dots, R_{n,m_n}$ is

$$\begin{aligned} & f(x_{n(1)}, r_{n,2}, \dots, r_{n,m_n}) \\ &= m_n! p^{m_n} x_{n(1)}^{-p-1} (r_{n,2} x_{n(1)})^{-p-1} (r_{n,3} r_{n,2} x_{n(1)})^{-p-1} \\ & \quad \cdot (r_{n,4} r_{n,3} r_{n,2} x_{n(1)})^{-p-1} \cdots (r_{n,m_n} \cdots r_{n,3} r_{n,2} x_{n(1)})^{-p-1} \\ & \quad \cdot x_{n(1)}^{m_n-1} r_{n,2}^{m_n-2} r_{n,3}^{m_n-3} r_{n,4}^{m_n-4} \cdots r_{n,m_n-1} \\ & \quad \cdot I(1 < x_{n(1)} < r_{n,2} x_{n(1)} < r_{n,3} r_{n,2} x_{n(1)} < \cdots < r_{n,m_n} \cdots r_{n,3} r_{n,2} x_{n(1)}). \end{aligned}$$

This is equivalent to

$$\begin{aligned} & f(x_{n(1)}, r_{n,2}, \dots, r_{n,m_n}) \\ &= m_n! p^{m_n} x_{n(1)}^{(-p-1)m_n+m_n-1} r_{n,2}^{(-p-1)(m_n-1)+m_n-2} \\ & \quad \cdot r_{n,3}^{(-p-1)(m_n-2)+m_n-3} r_{n,4}^{(-p-1)(m_n-3)+m_n-4} \cdots r_{n,m_n-1}^{(-p-1)2+1} r_{n,m_n}^{-p-1} \\ & \quad \cdot I(x_{n(1)} > 1) I(r_{n,2} > 1) I(r_{n,3} > 1) \cdots I(r_{n,m_n} > 1) \end{aligned}$$

which simplifies to

$$\begin{aligned} & f(x_{n(1)}, r_{n,2}, \dots, r_{n,m_n}) \\ &= m_n! p^{m_n} x_{n(1)}^{-pm_n-1} r_{n,2}^{-pm_n+p-1} r_{n,3}^{-pm_n+2p-1} r_{n,4}^{-pm_n+3p-1} \\ & \quad \cdots r_{n,m_n-1}^{-2p-1} r_{n,m_n}^{-p-1} I(x_{n(1)} > 1) I(r_{n,2} > 1) I(r_{n,3} > 1) \cdots I(r_{n,m_n} > 1). \end{aligned}$$

Integrating out the unnecessary random variable, $X_{n(1)}$, we have the joint density of our ratios of adjacent order statistics

$$\begin{aligned} & f(r_{n,2}, \dots, r_{n,m_n}) \\ &= m_n! p^{m_n} \int_1^\infty x_{n(1)}^{-pm_n-1} dx_{n(1)} \\ & \quad \cdot r_{n,2}^{-pm_n+p-1} r_{n,3}^{-pm_n+2p-1} r_{n,4}^{-pm_n+3p-1} \cdots r_{n,m_n-1}^{-2p-1} r_{n,m_n}^{-p-1} \\ &= (m_n - 1)! p^{m_n-1} r_{n,2}^{-pm_n+p-1} r_{n,3}^{-pm_n+2p-1} r_{n,4}^{-pm_n+3p-1} \cdots r_{n,m_n-1}^{-2p-1} r_{n,m_n}^{-p-1} \\ & \quad \cdot I(r_{n,2} > 1) I(r_{n,3} > 1) \cdots I(r_{n,m_n} > 1). \end{aligned}$$

Once again the random variables $R_{n,2}, \dots, R_{n,m_n}$ are independent, but not identically distributed. If $M_n = \max_{2 \leq k \leq m_n} R_{n,k}$, then

$$\begin{aligned} F_{M_n}(a) &= P\{M_n \leq a\} \\ &= P\{R_{n,2} \leq a\} P\{R_{n,3} \leq a\} P\{R_{n,4} \leq a\} \cdots P\{R_{n,m_n} \leq a\} \end{aligned}$$

$$\begin{aligned}
&= (m_n - 1)! p^{m_n - 1} \int_1^a r_{n,2}^{-pm_n + p - 1} dr_{n,2} \int_1^a r_{n,3}^{-pm_n + 2p - 1} dr_{n,3} \int_1^a r_{n,4}^{-pm_n + 3p - 1} dr_{n,4} \\
&\quad \cdots \int_1^a r_{n,m_n - 1}^{-2p - 1} dr_{n,m_n - 1} \int_1^a r_{n,m_n}^{-p - 1} dr_{n,m_n} \\
&= \left(1 - a^{-p(m_n - 1)}\right) \left(1 - a^{-p(m_n - 2)}\right) \cdots \left(1 - a^{-2p}\right) \left(1 - a^{-p}\right).
\end{aligned}$$

By selecting different values for p we can obtain various limit theorems. The next proof is the same as the last and will be omitted.

Theorem 2. *If $X_{n,1}, \dots, X_{n,m_n}$ are i.i.d. Pareto random variables with $p = 1$ and $\sup_{n \geq 1} m_n < \infty$, then for all $\alpha > 0$ we have*

$$\lim_{N \rightarrow \infty} \frac{\sum_{n=1}^N \frac{(\lg n)^{\alpha - 2}}{n} M_n}{(\lg N)^\alpha} = \frac{1}{\alpha} \quad \text{almost surely.}$$

If p is larger than one, then we obtain classic strong laws, where

$$f_{M_n}(a) = pa^{-p-1} + 2pa^{-2p-1} \dots$$

Theorem 3. *If $X_{n,1}, \dots, X_{n,m_n}$ are i.i.d. Pareto random variables with $p > 1$, then for a fixed sample size, $m_n = m$, we have*

$$\lim_{N \rightarrow \infty} \frac{\sum_{n=1}^N M_n}{N} = \frac{p}{p-1} + \frac{2p}{2p-1} + \dots \quad \text{almost surely.}$$

Proof. This follows from the fact that

$$E(M_n) = \int_1^\infty a f_{M_n}(a) da. \quad \square$$

4. Other Distributions

Note that if the random variable X has the density $f_X(x) = px^{-p-1}I(x \geq 1)$ and if $Y = 1/X$, then Y has the density $f_Y(y) = py^{p-1}I(0 < y < 1)$. That allows these results to hold for a sample from this particular Beta distribution. The reason is that the order statistics from these two

distributions are inverted

$$\frac{Y_{n(2)}}{Y_{n(1)}} = \frac{X_{n(m_n-1)}^{-1}}{X_{n(m_n)}^{-1}} = \frac{X_{n(m_n)}}{X_{n(m_n-1)}}$$

and

$$\frac{Y_{n(3)}}{Y_{n(2)}} = \frac{X_{n(m_n-2)}^{-1}}{X_{n(m_n-1)}^{-1}} = \frac{X_{n(m_n-1)}}{X_{n(m_n-2)}}$$

and so on. Hence

$$\max\left\{\frac{Y_{n(2)}}{Y_{n(1)}}, \frac{Y_{n(3)}}{Y_{n(2)}}, \frac{Y_{n(4)}}{Y_{n(3)}}, \dots, \frac{Y_{n(m_n)}}{Y_{n(m_n-1)}}\right\}$$

equals

$$\max\left\{\frac{X_{n(m_n)}}{X_{n(m_n-1)}}, \frac{X_{n(m_n-1)}}{X_{n(m_n-2)}}, \frac{X_{n(m_n-2)}}{X_{n(m_n-3)}}, \dots, \frac{X_{n(2)}}{X_{n(1)}}\right\}.$$

We can do this for many other distributions as well. Let Y_1, \dots, Y_{m_n} be i.i.d. exponential random variables with parameter β . Thus their underlying density is $f_Y(y) = \beta^{-1}e^{-y/\beta}I(y > 0)$. From Theorem 1, with $\theta_n = 1$, we have the result holding for uniform $(0,1)$ random variables and since we can obtain our random variable Y via the decreasing transformation $y = -\beta \ln x$, these strong laws also hold for exponential random variables. Likewise a strictly increasing transformation works just as well. The point is that any increasing or decreasing transformation of either the uniforms or these Paretos will produce similar results.

References

1. A. Adler, Exact laws for sums of order statistics from the pareto distribution, *Bull. Inst. Math. Acad. Sinica*, **31**(2003), N. 3, 181-193.
2. A. Adler, Exact laws for sums of ratios of order statistics from the pareto distribution, *Central European Journal of Math.*, **4** (2006), 1-4.
3. A. Adler, Laws of large numbers for ratios of uniform random variables, *De Gruyter Open Math.*, **13** (2015), 571-576.
4. A. Adler, Strong laws for ratios of order statistics from exponentials, *Bull. Inst. Math. Acad. Sinica*, (N.S.) **10**(1) (2015), 101-111.

5. Y. S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, 3rd ed., Springer-Verlag, New York, 1997.
6. Y. Miao, R. J. Wang and A. Adler, Limit theorems for order statistics from exponentials, *Stat. Probab. Letters*, **110** (2016), 51-57.
7. Y. Miao, Y. Sun, R. Wang and M. Dong, Various limit theorems for ratios from the uniform distribution, *De Gruyter Open Math.*, **14** (2016), 393-403.
8. Y. Zhang and X. Ding, Limit properties for ratios of order statistics from exponentials, *Journal of Inequality and Applications*, **11** (2017), doi:10.11861s13660-016-1287-6.