Finiteness of Record values and Alternative Asymptotic Theory of Records with Atom Endpoints

Gane Samb Lo [1]**, Harouna Sangaré** [2,3,*] **, Mamadou Cherif Moctar Traoré** [3] **and Mohammad Ahsanullah** [4]

[1] LERSTAD, Gaston Berger University, Saint-Louis, SENEGAL
Evanston Drive,NW, Calgary, Canada, T3P 0J9
Associate Researcher, LSTA, Pierre et Marie University, Paris, FRANCE
Professor, African University of Sciences and Technology, Abuja, NIGERIA
[2] DER MI, Faculté des Sciences et Techniques, USTT-B, MALI.
[3] LERSTAD, Gaston Berger University, Saint-Louis, SENEGAL.
[4] Department of Management Sciences, Rider University, Lawrenceville, New Jersey, USA.

Abstract. Asymptotic theories on record values and times, including central limit theorems, make sense only if the sequence of records values (and of record times) is infinite. If not, such theories could not even be an option. In this paper, we give necessary and/or sufficient conditions for the finiteness of the number of records. We prove, for example for *iid* real valued random variable, that strong upper record values are finite if and only if the upper endpoint is finite and is an atom of the common cumulative distribution function. (to be continued on page 2372).

* Corresponding author: Harouna Sangaré (harounasangare@fst-usttb-edu.ml, harouna.sangare@mesrs.ml, sangare.harouna@ugb.edu.sn)
Gane Samb Lo: gane-samb-lo@ugb.edu.sn
Mamadou Cherif Moctar Traoré: traore.cherif-mamadou-moctar@ugb.edu.sn
Mohammad Ahsanullah : ahsan@rider.edu

Full Abstract Asymptotic theories on record values and times, including central limit theorems, make sense only if the sequence of records values (and of record times) is infinite. If not, such theories could not even be an option. In this paper, we give necessary and/or sufficient conditions for the finiteness of the number of records. We prove, for example for *iid* real valued random variable, that strong upper record values are finite if and only if the upper endpoint is finite and is an atom of the common cumulative distribution function. The only asymptotic study left to us concerns the infinite sequence of hitting times of that upper endpoints, which by the way, is the sequence of weak record times. The asymptotic characterizations are made using negative binomial random variables and the dimensional multinomial random variables. Asymptotic comparison in terms of consistency bounds and confidence intervals on the different sequences of hitting times are provide. The example of a binomial random variable is given.

Résumé. Les théories asymptotiques sur les valeurs et les temps record, y compris les théorèmes centraux limites, n'ont de sens que si la suite de valeurs record (et de temps record) est infinie. Sinon, telles théories ne pourraient même pas être envisageables. Dans ce papier, nous donnons des conditions nécessaires et/ou suffisantes pour la finitude du nombre des records. Nous prouvons, par exemple pour les variables aléatoires iid à valeurs réelles, que les fortes valeurs record supérieures sont finies si et seulement si l'extrémité supérieure est finie et est un atome de la fonction de distribution cumulative commune. La seule étude asymptotique qui nous reste est celle de la suite infinie des temps d'atteinte de ces extrémités supérieures, qui sont d'ailleurs la suite de faibles temps record. les caractérisations asymptotiques sont effectuées à l'aide de variables aléatoires binomiales négatives et les variables aléatoires multinomiales dimensionnelles. Des comparaisons asymptotiques en termes de limites uniformes et d'intervalles de confiance sur les différentes suites de temps d'atteinte sont fournies. L'exemple d'une variable aléatoire binomiale est donné.

The authors.

Gane Samb Lo, Ph.D., is full professor at Gaston Berger University, Saint-Louis, SENEGAL and at the African University of Sciences and Technology, Abuja, Nigeria

Harouna Sangaré, Ph.D., is associate professor at the Faculté des Sciences et Techniques of the Universty of Sciences, Techniques and Technologies (USTTB), bamako, Mali.

Mamadou Cherif Moctar Traoré. M.Sc., is completing his Ph.D., under the supervision of the first author at Gaston Berger University, Saint-Louis, SENEGAL.

Mohammad Ahsanullah, Ph.D., is Emeritus professor at the Department of Management Sciences, Rider University, Lawrenceville, New Jersey, USA.

1. Introduction

A considerable number of asymptotic results are available in the literature concerning infinite sequences of record values of sequence of real-valued random variables defined on the same probability space (see for example Tata(1969), Ahsanullah and Nevzorov (2015), Resnick (1987), Nevzorev (2001), Lo and Ahsanullah (2018), etc., to cite a few). However if the upper-endpoint x_0 of the common cumulative distribution function (*cdf*) of a sequence of independent of identically distributed (*iid*) copies of X is an atom of the common probability law $\mathbb{P}_X$, that is $\mathbb{P}_X(\{x_0\})$ is not zero, there will not be any further record value once x_0 is hit for the first time. In such situations, all the available asymptotic theories become irrelevant. The paper shows the facts already described above and proposes asymptotic results on the sequences of hitting times of the upper end-point or the lower endpoints.

Let $X_1, X_2, \cdots$ be a sequence of real-valued random variables defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let us defined the sequence of record values:

$X^{(1)} = X_1$, $U(1) = 1$, and upon the existence of $Y^{(n)} = X_{U(n)}$

$$X^{(n+1)} = \begin{cases} \inf\{j > U(n), X_j > Y^{(n)}\}, \\ \text{Not defined if } \inf\{j > U(n), X_j > Y^{(n)}\} = +\infty. \end{cases}$$

The sequence $\{U(n), 1 \leq n \leq M\}$ is that of the occurence times of the record values, where M can be infinite or finite, constant or random.

If there is no further record value after $X^{(n)}$, we denote $U(n+1) = +\infty$.

In this note, we begin by a general result on the finiteness or infiniteness of the total number of records regardless the dependence structure of $(X_n)_{n \geq 1}$ in Section 2. After the justification of the finiteness of upper record if $uep(F)$ is an atom and the finiteness of lower records if $lep(F)$ is an atom, we study the infinite sequences of hitting times of one of the endpoints in Section 3. We provide central limit theorems, laws of the iterated logarithm and Berry-Essence bounds for the sequence of hitting times $(N_{i,k})_{k \geq 1}$ of $lep(F)$ ($i = 1$) and of $uep(F)$ ($i = 2$) (Theorem 2). These results are based on probability laws of negative binomial random variable.

Next, in the same section, we take a multinomial approach with three outcomes E_1 (lep(F) is hit) of probability p_1, E_2 (uep(F) is hit) of probability p_2 and E_3 (neither lep(F) nor uep(F) is hit) of probability $p_3 = 1 - p_1 - p_2$. We study the asymptotic law of the mutinomial random vector $(M_{1,n}, M_{2,n}, M_{3,n})^t$ (Proposition 4), and draw asymptotic sub-results on the difference $M_{1,n} - M_{2,n}$ and on the ratio $M_{1,n}/M_{2,n}$ in Corollary 1.

In summary, for $p_1 - p_2 > 0$ for example, the difference $M_{1,n} - M_{2,n}$ goes to infinity with the rate δn, where $0 < \delta < p_1 - p_2$. In section 4, we give an illustration in the

case where X is a binomial random variable $\beta(r,\alpha)$, $r \geq 1$, $\alpha \in]0,1[$ and provide confidence interval for $M_{1,n}/M_{2,n}$ in the case where $\alpha = 1/2$. We finish the paper by a conclusive section.

2. Finiteness or Infiniteness of the total number of records

Let us begin by a general law.

Proposition 1. *For each $k \geq 1$, set*

$$X_k^\star = \sup_{h>k} X_h$$

and denote

$$D_- = \{(x,y) \in \mathbb{R}^2, \ x \leq y\}.$$

We have for any $n \geq 2$

$$\mathbb{P}(U(n+1) = +\infty) = \sum_{k \geq n} \mathbb{P}_{(X_k^\star, X_k)}(D_-)\mathbb{P}(U(n) = k).$$

Proof. Conditioning on $(U(n) = k)$, $(U(n+1) = +\infty)$ means that all the X_h, $h > k$, are less than X_k. The proof is ended by the remark

$$\mathbb{P}(\max_{h>k} X_h \leq X_k) = \mathbb{P}(X_k^\star \leq X_k) = \mathbb{P}((X_k^\star, X_k) \in D_-) = \mathbb{P}_{(X_k^\star, X_k)}(D_-).\blacksquare$$

Let us give an application of Proposition 1 in the independent case.

Proposition 2. *Suppose that $X_1, X_2, \ldots$ are independent random variables with respective cumulative distribution functions (cdf) F_j, $j \geq 1$. Then, whenever $U(n)$ is finite, we have*

$$\mathbb{P}(U(n+1) = +\infty) = \sum_{k \geq n} \left(\int_{\mathbb{R}} \left(\prod_{j>k} F_j(x) \right) d\mathbb{P}_{X_k}(x) \right) \mathbb{P}(U(n) = k).$$

Proof. Here $X_k^\star$ and X_k are independent and we have

$$\mathbb{P}(U(n+1) = +\infty) = \sum_{k \geq n} \mathbb{P}_{X_k^\star} \otimes \mathbb{P}_{X_k}(D_-)\mathbb{P}(U(n) = k).$$

Let us use Fubini's theorem to have

$$\mathbb{P}_{X_k^\star} \otimes \mathbb{P}_{X_k}(D_-)$$
$$= \int_{\mathbb{R}} d\mathbb{P}_{X_k}(x) \int_{\mathbb{R}} 1_{D_-}(x,y) d\mathbb{P}_{X_k^\star}$$
$$= \int_{\mathbb{R}} \mathbb{P}(X_k^\star \leq x) d\mathbb{P}_{X_k}(x)$$
$$= \int_{\mathbb{R}} \left(\prod_{j>k} F_j(x) \right) d\mathbb{P}_{X_k}(x).$$

We get the announced result by combining the above lines.

Now let us see what happens if the sequence is stationary, that is $F_j = F$ for all $j \geq 1$. We define the lower and the upper endpoints (lep and uep) of F by

$$lep(F) = \inf\{x \in \mathbb{R}, \ F(x) > 1\} \ and \ uep(F) = \sup\{x \in \mathbb{R}, \ F(x) < 1\}.$$

We have

$$\int_{\mathbb{R}} \left(\prod_{j>k} F_j(x) \right) d\mathbb{P}_{X_k}(x) = \int_{-\infty}^{uep(F)} F(x)^{+\infty} dF(x).$$

But $F(x)^{+\infty} = 0$ unless $x = uep(F)$. This gives

$$\int_{\mathbb{R}} \left(\prod_{j>k} F_j(x) \right) d\mathbb{P}_{X_k}(x) = \int_{-\infty}^{uep(F)} 1_{\{uep(F)\}} dF(x) = \mathbb{P}(X = uep(F)).$$

We conclude that

$$\mathbb{P}(U(n+1) = +\infty) = \sum_{k \geq n} \mathbb{P}(X = uep(F))\mathbb{P}(U(n) = k) = \mathbb{P}(X = uep(F))$$

which leads to the simple result:

Proposition 3. *Suppose that $X_1, X_2, \ldots$ are independent and identically distributed random variables with common cdf F and let $uep(F)$ denote the upper endpoint of F. Then*

$$\mathbb{P}(U(n+1) = +\infty) = \mathbb{P}(X = uep(F)).$$

As a consequence, the sequence of record values (and of record times) is finite if and only if $uep(F)$ is finite and is an atom of F, that is $\mathbb{P}_X(uep(F)) > 0$.

Consequences. The number of time records *a.s.* is infinite in the following cases.

(1) $uep(F) = +\infty$.

(2) $uep(F) < +\infty$ but $\mathbb{P}(X = uep(F)) = 0$. **Example** : $X \sim \mathcal{U}(0,1)$.

The number of time records may be finite in the following cases.

(1) X is discrete and takes a finite number of points. Such a remark has been made by Shorrock (1987).

(2) X is discrete, takes an infinite number of values such that the strict values set $\mathcal{V}_X$ of X has a maximum value. We mean by strict values set, the set of points taken by X with a non-zero probability.

3. Infinite number of hitting times for the extreme endpoints

In this note, we focus on X, X_1, X_2, ... are *iid* random variables with *cdf* F such that $uep(F)$ is finite and is an atom of F, that is

$$\mathbb{P}(X = uep(F)) = p_2 \in \,]0,1[\,. \tag{1}$$

In that context, let us see right now that $uep(F)$ will be hit infinitely many times, that is

$$\{j \geq 1, \ X_j = uep(F)\}$$

forms an infinite sequence of random variables

$$(N_{k,2})_{k \geq 1} = \{N_{1,2}, N_{2,2}, \cdots\}, \tag{2}$$

Indeed, by denoting the event $uep(F)$ will not hit by any X_j, $j \geq 1$ by A_n, we have for $n = 1$

$$\mathbb{P}(A_1) = \mathbb{P}\left(\bigcap_{j \geq 1}(X_j \neq uep(F))\right) = (1 - p_2)^{+\infty} = 0,$$

and for any $n \geq 1$,

$$\begin{aligned}
\mathbb{P}(A_n) &= \sum_{k \geq 1} \mathbb{P}(A_n/(N_{n-1,2} = k))\mathbb{P}(N_{n-1,2} = k) \\
&= \mathbb{P}\left(\bigcap_{j \geq 1}(X_j \neq uep(F))\right)\mathbb{P}(N_{n-1,2} = k) \\
&= (1 - p_2)^{+\infty}\,\mathbb{P}(N_{n-1,2} = k) = 0. \ \square
\end{aligned}$$

G.S. Lo, H. Sangaré, C. M. M. Traoré and M. Ahsanullah, Afrika Statistika, Vol. 15 (3),
2020, pages 2371 - 2385. Finiteness of Record values and Alternative Asymptotic Theory
of Records with Atom Endpoints.   2377

As well, if

$$\mathbb{P}\left(X = lep\left(F\right)\right) = p_1 \in \left]0, 1\right[. \tag{3}$$

The sequence of hitting times of $lep(F)$ is infinite and is denoted as

$$(N_{k,1})_{k \geq 1} = \{N_{1,1}, N_{2,1}, \cdots\}.$$

Throughout the paper, we suppose that $0 < p_1 + p_2 < 1$, otherwise X would be a
Bernoulli random variable and the study would be quite simple. So, if both (1) and
(2) holds, we may define

$$(N_{k,3})_{k \geq 1} = \{N_{1,3}, N_{2,3}, \cdots\},$$

as the random times in which neither endpoint is hit, that is

$$\{j \geq 1, \ X_j \neq lep(F) \ and \ X_j \neq uep(F)\}.$$

We wish to describe the asymptotic theory of such sequences when they are de-
fined. That theory reduces to studying sequences of negative binomial random
variables. Let us make that first recall.

Theorem 1. *Let X, X_1, X_2, . . . be iid random variables with common cdf F.*

*(a) If $uep(F) \in \mathbb{R}$ and (1) holds, the sequence $(N_{k,2})_{k \geq 1}$ is an infinite sequence and for
all $k \geq 1$, $N_{k,2}$ follows a negative binomial law of parameters and k and $\bar{p}_2$*

$$N_{k,2} \sim \beta N\left(k, p_2\right).$$

*(b) If $lep(F) \in \mathbb{R}$ and (3) holds. Then the sequence $(N_{k,1})_{k \geq 1}$ is infinite and for all $k \geq 1$,
$N_{k,1}$ follows a negative binomial law of parameters and $\overline{k}$ and $\bar{p}_1$*

$$N_{k,1} \sim \beta N\left(k, p_1\right).$$

(c) (1) and (2) hold and $0 < p_1 + p_2 < 1$, the sequence of $(N_{k,3})_{k \geq 1}$ is infinite and

$$\forall k \geq 1, N_{k,3} \sim \beta \mathcal{N}\left(k, p_3\right).$$

Proof. It is enough to prove this for one case. For example, if (1) holds, $Z_{1,2} = N_{1,2}$
follows a geometric law of parameter p_2, that is $Z_{1,2} \sim \mathcal{G}\left(p_2\right)$ and for all $h \geq 1$,

$$\mathbb{P}\left(Z_{1,2} = h\right) = \bar{p}_2^{h-1}p_2,$$

with

$$\overline{p}_i = 1 - p_i, \ i \in \{1,2,3\}.$$

Once $Z_{1,2}$ is observed, the next hitting time is achieved after $Z_{2,2}$ trials and

$$N_{2,2} = Z_{1,2} + Z_{2,2}$$

with $Z_{1,2}$ independent of $Z_{2,2}$. By induction, we get $Z_{1,2}, Z_{2,2}, \cdots, Z_{k,2}, \cdots$ independent such that each $Z_{k,2} \sim \mathcal{G}(p_2)$ so that for $k \geq 1$,

$$N_{k,2} = Z_{1,2} + \cdots + Z_{k,2}.$$

This proves that $N_{k,2} \sim \beta N(k, p_2)$. $\square$

From the previous proof, we saw that we have, for $i \in \{1,2,3\}$, that there exists a sequence of independent and geometric random variables $(Z_{k,i})_{k \geq 1}$ such that for all $k \geq 1$.

$$N_{k,i} = Z_{1,i} + \cdots + Z_{k,i}, \ the \ Z'_{\circ,i}s \ are \ independent, \ Z_{\circ,i} \sim \mathcal{G}(p_i).$$

From there, the classical limit theorems for sums of *iid* random variables apply as follows.

We denote, for $i \in \{1,2,3\}$

$$\nu_i = \mathbb{E}(Z_{1,i}) = 1/p_i,$$

$$\sigma_i^2 = \mathbb{V}ar(Z_{1,i}) = \overline{p}_i/(p_i^2),$$

$$\gamma_i = \mathbb{E}(Z_{1,i} - \mathbb{E}Z_{1,i})^3 = \frac{\overline{p}_i(2 - p_i)}{p_i^3}.$$

We have

Theorem 2. *Upon* (1) *holds for* $i = 1$, (3) *holds for* $i = 2$ *and both* (1) *and* (3) *hold for* $i = 3$ *with* $p^{(3)} < 1$, *we have the following laws for* $i \in \{1,2,3\}$.

(a) **Central limit theorem (CLT).**

$$N_k^{(i,\star)} = \frac{N_k^{(i)} - k\nu_i}{\sigma_i\sqrt{k}} \rightsquigarrow \mathcal{N}(0,1)$$

(b) Iterated logarithm law (LIL).

$$\lim_{k \to +\infty} \frac{N_k^{(i)} - k\nu_i}{\sqrt{2k\sigma_i \log\log k}} = +1$$

and

$$\limsup_{k \to -\infty} \frac{N_k^{(i)} - k\nu_i}{\sqrt{2k\sigma_i \log\log k}} = -1$$

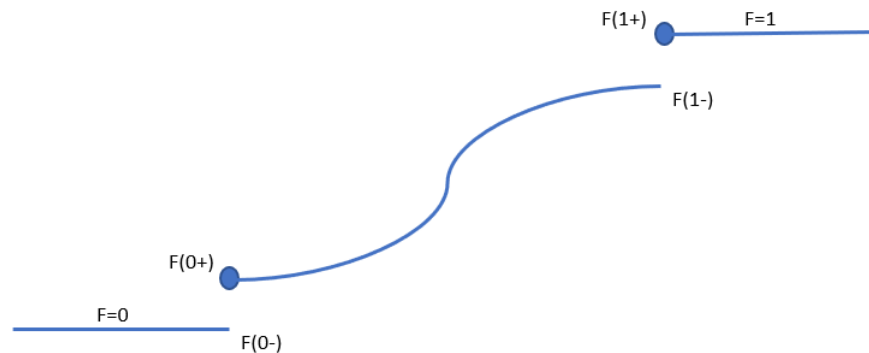*(c) Berry-Essen approximation (BEA).*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}\left(N_k^{(i,\star)} \leq x \right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt \right| \leq \frac{36\gamma_i}{\sqrt{k}}.$$

The limit theorems (CLT), (LIL) and (BEA) are to be found in graduate probability textbooks like Loeve (1997), Chung (1974), Gutt (2005), etc., in particular in Lo (2018) (Theorem 20 page 237, Theorem 22 page 272, Theorem 21 page 253).

If both endpoints are finite and atoms for the probability law, the *cdf* may be represented as follows

Fig. 1. The two finite endpoints are both atoms

In such cases, it would be interesting to compare the number of hitting times of one endpoint with the same number of the other endpoints or with number of hitting times of the non-endpoint zone.

Let us proceed by a multinomial approach. For each observation time $i \geq 1$, we simultaneously check which of the three events $A_i = (X_i = lep(F))$,

$B_i = (X_i \notin \{lep(F), uep(F)\})$ or $C_i = (X_i = uep(F))$ occurs.

At time $n \geq 1$, $M_{n,i}$ is the number of occurrences of the A_i's, of the C_i's and finally of the B_i's. So, we have that $M_{n,1}$ is number of hitting of $lep(F)$, $M_{n,2}$ the number of hitting of $uep(F)$ and $M_{n,3}$ the number of times neither $lep(F)$ nor $uep(F)$ are hit.

The vector $M_n = (N_{n,1}, M_{n,2}, M_{n,3})$ follows a three dimensional multinomial law of parameters (n, p), with $p = (p_1, p_2, p_3)$. We suppose that $p_i > 0$, $i \in \{1, 2, 3\}$ and $p_1 + p_2 + p_3 = 1$.

We focus on compare $M_{n,1}$ and $M_{n,2}$ for large values of n, the results being extended to any pair of $\{(M_{n,1})_{n \geq 1}, (M_{n,2})_{n \geq 1}, (M_{n,3})_{n \geq 1}\}$.

Let us use the classical result of the asymptotic law of $(M_n)_{n \geq 0}$.

Proposition 4. *Given the context described above and the assumptions set above, we have*

$$M_n^{\star} = \left(\frac{M_{n,1} - np_1}{\sqrt{np_1}}, \frac{M_{n,2} - np_2}{\sqrt{np_2}}, \frac{M_{n,3} - np_3}{\sqrt{np_3}} \right)^t$$

converges to a centered Gaussian vector $Z = (Z_1, Z_2, Z_3)^t$ with variance-covariance $\Sigma = (\sigma_{ij})_{1 \leq i,j \leq 3}$:

$$\begin{cases} \sigma_{ii} = 1 - p_i \\ \sigma_{ij} = -\sqrt{p_i p_j}, \ i \neq j. \end{cases}$$

We infer the following laws.

Corollary 1. *Given the notation above, we have*
(a)

$$\sqrt{n} \left(\frac{M_{n,1}}{M_{n,2}} - \frac{p_1}{p_2} \right) \sim \mathcal{N}\left(0, \gamma_{1,2}^2\right)$$

with

$$\gamma_{1,2}^2 = (p_1/p_2)(p_2^2 \overline{p}_1 + p_1^2 \overline{p}_2 + 2(p_1 p_2)^{3/2}).$$

(b)

$$\sqrt{n} \left(\frac{M_{n,1} - M_{n,2}}{n} - (p_1 - p_2) \right) \sim \mathcal{N}\left(0, \delta_{1,2}^2\right)$$

with

$$\delta_{1,2}^2 = p_1 \overline{p}_1 + p_2 \overline{p}_2 + 2(p_1 p_2).$$

Before we provide the proof of Corollary 1, we give some of its important consequences. We are going to get the following. If the hitting probability of one of the endpoints is greater than the other by $\Delta p = |p_1 - p_2| > 0$, for any $\beta in]0, \Delta p[$, its number of hitting times is greater than the other counterpart of more that $n\beta$, that is (if $p_1 > p_2$ for example)

$$\liminf_{n \to +\infty} \mathbb{P}(M_{n,1} - M_{n,2} \ge n\beta) = 100\%.$$

To see this, suppose that $\Delta = p_1 - p_2 > 0$ and $0 < \beta < \Delta$ and set $I(n) = (M_{n,1} - M_{n,1} \ge n\beta)$, $n \ge 1$. By part (b) of Corollary 1, and by G the *cdf* of absolute value the standard Gaussian random variable and by denoting

$$C_n = \left| \sqrt{n} \left(\frac{M_{n,1} - M_{n,2}}{n} - (p_1 - p_2) \right) \right|.$$

We have for any $t \in \mathbb{R}$, $\mathbb{P}(C_n \ge t) - (1 - G(t)) \to 0$ and by a classical result in Weak convergence (when the *limit cdf* is continuous, see for example Lo *et al.*(2016) [page 107, Chapter 4, Point (5)], we get

$$\delta_n = \sup_{t \in \mathbb{R}} |\mathbb{P}(C_n \ge t) - (1 - G(t))| \to 0.$$

Let us apply this to

$$
\begin{aligned}
\mathbb{P}(I(n)) &= \mathbb{P}\left(\frac{M_{n,1} - M_{n,2}}{n} \le \Delta p \right) \\
&= \mathbb{P}\left(\sqrt{n} \left(\frac{M_{n,1} - M_{n,2}}{n} - (p_1 - p_2) \right) \le \sqrt{n}(\beta - \Delta p) \right).
\end{aligned}
$$

We remark that for each $n \ge 1$, $-x_n = \sqrt{n}(\beta - \Delta p)$ is negative and $x_n \to +\infty$ as $n \to +\infty$. So we have

$$
\begin{aligned}
\mathbb{P}(I(n)^c) &\le \mathbb{P}\left(\left| \sqrt{n} \left(\frac{M_{n,1} - M_{n,2}}{n} - (p_1 - p_2) \right) \right| \ge x_n \right) \\
&\le |\mathbb{P}(C_n \ge x_n) - (1 - G(x_n))| + (1 - G(x_n)) \\
&\le \delta_n + (1 - G(x_n)) \to 0,
\end{aligned}
$$

as $n \to +\infty$. We have proved that

G.S. Lo, H. Sangaré, C. M. M. Traoré and M. Ahsanullah, Afrika Statistika, Vol. 15 (3),
2020, pages 2371 - 2385. Finiteness of Record values and Alternative Asymptotic Theory
of Records with Atom Endpoints.   2382

$$\liminf_{n \to +\infty} \mathbb{P}(I(n)) = 100\%. \ \square$$

Proof of Corollary 1. We use the Skorohod-Wichura theorem (See Skorohod (1956)
and Wichura (1970) and Lo (2019) for a brief proof) and place ourselves on a prob-
ability space holding a sequence $\left(\overline{N}_n^\star\right)_{n \geq 1}$ and $\overline{Z}$ such that

$$\forall n \geq 1, \overline{N}_n^\star =_d M_n^\star, \ Z =_d \overline{Z}$$

and

$$\overline{N}_n^\star \longrightarrow_{\mathbb{P}} \overline{Z}.$$

So, we may and do take $M_n^\star = \overline{N}_n^\star$ and $Z = \overline{Z}$. We have

$$\frac{M_{n,1}}{M_{n,2}} = \frac{\sqrt{np_1}\,(Z_1 + o_p(1)) + np_1}{\sqrt{np_2}\,(Z_2 + o_p(1)) + np_2}.$$

From there, as proved in the Appendix,

$$\sqrt{n}\left(\frac{M_{n,1}}{M_{n,2}} - \frac{p_1}{p_2}\right) = \left(\frac{p_1}{p_2}\right)^{1/2}(p_2 Z_1 - p_1 Z_2)$$
$$\sim \mathcal{N}\left(0, \gamma_{1,2}^2\right),$$

with

$$\gamma_{1,2}^2 = \mathbb{V}ar\left((p_1 p_2)^{1/2}(p_2 Z_1 - p_1 Z_2)\right)$$
$$= (p_1/p_2)\left[p_2^2 \mathbb{E}Z_1^2 + p_1^2 \mathbb{E}Z_2^2 - 2p_1 p_2 \mathbb{E}\left(Z_1 Z_2\right)\right].$$

Finally,

$$\gamma_{1,2}^2 = (p_1/p_2)\left[p_2^2\left(1 - p_1\right) + p_1^2\left(1 - p_2\right) + 2\left(p_1 p_2\right)^{3/2}\right].$$

We also have

$$M_{n,1} - M_{n,2} = \left(\sqrt{np_1}\,(Z_1 + o_p(1)) + np_1\right) - \left(\sqrt{np_2}\,(Z_2 + o_p(1)) + np_2\right)$$
$$= n\left(p_1 - p_2\right) + \sqrt{n}\left(\sqrt{p_1}\,(Z_1 + o_p(1)) - \sqrt{p_2}\,(Z_2 + o_p(1))\right).$$

So,

$$\sqrt{n}\left(\frac{M_{n,1}-M_{n,2}}{n}-(p_1-p_2)\right)=\sqrt{p_1}Z_1-\sqrt{p_2}Z_2+o_p(1).$$

Since

$$\begin{aligned}\delta_{1,2}^2 &= \mathbb{V}ar\left(\sqrt{p_1}Z_1-\sqrt{p_2}Z_2\right)\\ &= p_1\left(1-p_1\right)+p_2\left(1-p_2\right)+2p_1p_2.\end{aligned}$$

So,

$$\sqrt{n}\left(\frac{M_{n,1}-M_{n,2}}{n}-(p_1-p_2)\right)\sim\mathcal{N}\left(0,\delta_{1,2}^2\right).\ \square$$

Appendix

$$\begin{aligned}\frac{M_{n,1}}{M_{n,2}}-\frac{p_1}{p_2} &= \frac{\sqrt{np_1}\left(M_n^\star(1)+np_1\right)}{\sqrt{np_2}\left(M_n^\star(2)+np_2\right)}-\frac{p_1}{p_2}\\ &= \frac{p_2\sqrt{p_1p_2}\sqrt{n}M_n^\star(1)+np_1p_2-\left(\sqrt{n}p_1\sqrt{p_1}p_2M_n^\star(2)+np_1p_2\right)}{np_2\left(M_n^\star(2)/\left(\sqrt{np_2}\right)+1\right)}\\ &= \frac{\sqrt{p_1p_2}\left(p_2M_n^\star(1)-p_1M_n^\star(2)\right)}{p_2\sqrt{n}\left(1+M_n^\star(1)/\sqrt{np_2}\right)}\\ &= \frac{\sqrt{p_1p_2}\left(p_2\left(Z_1+o_P(1)\right)-p_1\left(Z_2+o_P(1)\right)\right)}{p_2\sqrt{n}\left(1+\left(Z_1+o_P(1)\right)/\sqrt{np_2}\right)}.\end{aligned}$$

Hence

$$\sqrt{n}\left(\frac{M_{n,1}}{M_{n,2}}-\frac{p_1}{p_2}\right)=\left(\frac{p_1}{p_2}\right)^{1/2}\left(p_2Z_1-p_1Z_2\right)+o_P(1).$$

So,

$$\sqrt{n}\left(\frac{M_{n,1}}{M_{n,2}}-\frac{p_1}{p_2}\right)\sim\mathcal{N}\left(0,\gamma_{1,2}^2\right)$$

with

$$\begin{aligned}\gamma_{1,2}^2 &= \mathbb{V}ar\left(\left(p_1/p_2\right)^{1/2}\left(p_2Z_1-p_1Z_2\right)\right)\\ &= (p_1/p_2)\left[p_2^2\mathbb{V}arZ_1+p_1^2\mathbb{V}arZ_2-2p_1p_2\mathbb{C}ov\left(Z_1Z_2\right)\right]\\ &= (p_1/p_2)\left[p_2^2\left(1-p_1\right)+p_1^2\left(1-p_2\right)+2\left(p_1p_2\right)^{3/2}\right].\end{aligned}$$

4. A simple application

Let X follows a binomial $\beta(r, \alpha)$, $r \geq 1$ and $\alpha \in]0, 1[$. The endpoints are $lep(F) = 0$ and $uep(F) = r$ which are atoms of probability $(1 - \alpha)^r$ and α^r.

If $1 - \alpha < \alpha$, i.e., $\alpha > 1/2$, we have, by Theorem , for any $0 < \delta < \alpha^r - (1 - \alpha)^r$.

$$\liminf \mathbb{P}(M_{n,2} - M_{n,1} > n\delta) = 100\%. \tag{4}$$

If $1 - \alpha > \alpha$, i.e., $\alpha < 1/2$, we have, by Part (b) of Corollary 1, for any $0 < \delta < (1 - \alpha)^r - \alpha^r$

$$\liminf \mathbb{P}(M_{n,1} - M_{n,2} > n\delta) = 100\%. \tag{5}$$

If $\alpha = 1/2$, $p_1 = p_2 = \alpha^3 = (1 - \alpha)^r = 2^{-r}$. By applying Part (a) of Corollary 1, we have

$$\gamma_{1,2} = 2\gamma^2 = 22^{-2r} = 2^{-2r+1}.$$

and

$$2^{r+1/2}\left((M_{n,1}/M_{n,2}) - 1\right) \rightsquigarrow \mathcal{N}(0, 1).$$

By defining z_u the critical points of a standard normal *cdf*, that is $\mathbb{P}(\mathcal{N}(0, 1) \leq z_u) = u$ for $u \in]0, 1[$, we get confidence interval :

$$\forall u \in]0, 1[, \ \mathbb{P}\left(1 - \frac{2^r z_{1-u/2}}{\sqrt{2n}} \leq \frac{M_{n,1}}{M_{n,2}} \leq 1 + \frac{2^r z_{1-u/2}}{\sqrt{2n}}\right) \approx 1 - u. \tag{6}$$

References

Ahsanullah M. and Nevzorov B.V.(2015) Records via Probability Theory. Atlantis Press.

Allan Gutt (2005). *Probability Theory : a graduate course.* Springer.

Kai Lai Chung (1974). *A Course in Probability Theory.* Academic Press. New-York.

Lo, G.S.(2018). *Mathematical Foundation of Probability Theory.* SPAS Books Series. Saint-Louis, Senegal - Calgary, Canada. Doi : http://dx.doi.org/10.16929/sbs/2016.0008. Arxiv : arxiv.org/pdf/1808.01713

Lo G.S. and Ahsanullah M. (2019). Asymptotic laws for upper and strong record values in the extreme domain of attraction and beyond. Arxiv : 1905.03380.

Lo, G.S., Niang E. B. and Chinewendu L.O (2020). A Detailed and Direct Proof of Skorohod-Wichura's Theorem.. In: Theory and Practice of Mathematics and Computer Science Vol. 1. [TPMCS-V1] (Ed. B. Chaouchi). **Chapter 8**, pp.91-107. ISBN: 978-93-90149-72-8 (print), 978-93-90149-22-3 (ebook). Doi: 10.9734/bpi/tpmcs/v1.

Lo, G.S.(2016). Weak Convergence (IA). Sequences of random vectors. SPAS Books Series. Saint-Louis, Senegal - Calgary, Canada. Doi : 10.16929/sbs/2016.0001. Arxiv : 1610.05415. ISBN : 978-2-9559183-1-9

Loève, Michel.(1997). *Probability Theory I.* Springer-Verlag, 4th Edition.

Nevzorov, V. B. (2001). *Records : Mathematical Theory.* Translation of Mathematical Monographs, Volume 194. American Mathematical Society. Providence, RI, USA

Resnick, S.I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer-Verbag, New-York. (MR0900810)

Shorrock R.W.(1972) On discrete Time Extremal processes. Advances in Applied Probability. Vol 6 (3), pp. 580-592

Skorohod A. V.(1956) Limit Theorems for Stochastic processes. Theor. probability Appl. 1, 261-290

Tata M.N. (1969) On utstanding values in a sequences of random variables. *Zeitschrift fuer Wehrsche. Verw. Gebiete*. Vol 12, pp. 9-20

Wichura M.J. (1970) On the Construction Of Almost sure Uniformly Convergent Random Variable with Given Weakly Convergent Images laws. *Ann. Math. Statist.*, 41 (1), pp. 284-291