A note on random permutations and extreme value distributions

By Pavle Mladenović

University of Belgrade, Faculty of Mathematics, Studentski trg 16, 11 000 Belgrade, Yugoslavia

(Communicated by Heisuke Hironaka, M. J. A., Oct. 15, 2002)

Abstract: Let Ω_n be the set of all permutations of the set $N_n = \{1, 2, \ldots, n\}$ and let us suppose that each permutation $\omega = (a_1, \ldots, a_n) \in \Omega_n$ has probability $1/n!$. For $\omega = (a_1, \ldots, a_n)$ let $X_{nj} = |a_j - a_{j+1}|$, $j \in N_n$, $a_{n+1} = a_1$, $M_n = \max\{X_{n1}, \ldots, X_{nn}\}$. We prove herein that the random variable M_n has asymptotically the Weibull distribution, and give some remarks on the domains of attraction of the Fréchet and Weibull extreme value distributions.

Key words: Random permutations; maximum of random sequence; Leadbetter's mixing condition; extreme value distributions; domains of attraction.

1. Introduction. Let Ω_n be the set of all permutations of the set $N_n = \{1, 2, \ldots, n\}$ and let us suppose that each permutation

$$\omega = (a_1, \ldots, a_n) \in \Omega_n$$

has probability $1/n!$. Random permutations have been very much studied and many asymptotic results as $n \to \infty$ have been obtained. For example, the number of cycles of a random permutation and the logarithm of the order of a random permutation are asymptotically normally distributed. See for example [3]. For $\omega = (a_1, \ldots, a_n)$ let us denote:

$$X_{nj}(\omega) = |a_j - a_{j+1}|, \quad j \in N_n,$$

where $a_{n+1} = a_1$ and

$$M_n = \max\{X_{n1}, \ldots, X_{nn}\}.$$

Then, $X_{n1}, \ldots, X_{nn}$ is a sequence of *dependent* random variables that satisfies condition of strict stationarity. It is easy to verify that for every $j \in N_n$, the marginal distribution of random variable X_{nj} is given by

$$P\{X_{nj} = k\} = \frac{2(n-k)}{n(n-1)}, \quad k \in \{1, 2, \ldots, n-1\}.$$

In this note we determine the limiting distribution of random variable M_n and give some remarks on the domains of attraction of the Fréchet and Weibull extreme value distributions.

Theorem 1. *For every real number x the following equality holds*:

$$\lim_{n\to\infty} P\left\{M_n \leqslant x\sqrt{n} + n\right\} = \begin{cases} e^{-x^2}, & \text{if } x < 0, \\ 1, & \text{if } x \geqslant 0. \end{cases}$$

2. Proof of Theorem 1. Let $X^*_{n1}, \ldots, X^*_{nn}$ be a sequence of n *independent* random variables which have the same distribution as random variables $X_{n1}, \ldots, X_{nn}$. Throughout this section we shall use the following notations: F_n – the common distribution function of random variables X_{nj} and X^*_{nj}, $j \in N_n$, and $M^*_n = \max\{X^*_{n1}, \ldots, X^*_{nn}\}$, $A_{nj} = \{X_{nj} > u_n\}$, $j \in N_n$.

Lemma 1 ([4], Theorem 1.5.1)**.** *Let (u_n) be a sequence of real numbers. Then, the equality*

$$\lim_{n\to\infty} n(1 - F_n(u_n)) = \tau$$

holds for $0 \leqslant \tau \leqslant +\infty$ if and only if

$$\lim_{n\to\infty} P\{M^*_n \leqslant u_n\} = e^{-\tau}.$$

Lemma 2. *The limiting distribution of random variable M^*_n is given by*

$$\lim_{n\to\infty} P\{M^*_n \leqslant x\sqrt{n} + n\} = \begin{cases} e^{-x^2}, & \text{if } x < 0, \\ 1, & \text{if } x \geqslant 0. \end{cases}$$

Proof. Let $F_n(x) = P\{X_{nj} \leqslant x\} = P\{X^*_{nj} \leqslant x\}$. It is easy to verify that for all positive integers $m \in \{1, 2, \ldots, n-1\}$ the following equalities hold:

$$F_n(m) = \frac{2}{n(n-1)}\left\{mn - \frac{m(m+1)}{2}\right\},$$

$$1 - F_n(m) = 1 - \frac{2m}{n-1} + \frac{m(m+1)}{n(n-1)}.$$

Let us denote $u_n = u_n(x) = x\sqrt{n} + n$. Then for $x < 0$ we obtain

1991 Mathematics Subject Classification. Primary 60G70; Secondary 05A05.

$$\lim_{n\to\infty} n(1-F_n(u_n)) = x^2,$$

and for $x \geqslant 0$ and every positive integer n we get $n(1-F_n(u_n)) = 0$. Consequently, the statement of Lemma 2 follows by Lemma 1. □

Lemma 3. *For $x < 0$ and $u_n = x\sqrt{n} + n$ the following asymptotic relations hold as $n \to \infty$:*

$$P(A_{nj}) \sim \frac{x^2}{n},$$

$$P(A_{n1}A_{n2}) \sim \frac{2(-x)^3}{3n^{3/2}},$$

$$P(A_{n1}A_{nj}) \sim \frac{x^4}{n^2}, \quad j \in \{3,\ldots,n-1\}.$$

$$P(A_{n1}A_{n2}A_{n3}) = O\Big(\frac{1}{n^2}\Big),$$

$$\cdots\cdots\cdots$$

Proof. Straightforward exercise. □

Lemma 4. *Let $x < 0$ and $u_n = x\sqrt{n} + n$. Then there exists a real constant $C_1(x)$, such that for every positive integer $k \leqslant n$ and all*

$$1 \leqslant j_1 < j_2 < \cdots < j_k \leqslant n$$

the following inequality holds:

$$\bigg|P\Big(\bigcap_{r=1}^{k} \overline{A}_{nj_r}\Big) - \prod_{r=1}^{k} P(\overline{A}_{nj_r})\bigg| \leqslant \frac{C_1(x)}{\sqrt{n}}.$$

Proof. The following equalities hold:

$$\begin{aligned}
&P\Big(\bigcap_{r=1}^{k} \overline{A}_{nj_r}\Big) - \prod_{r=1}^{k} P(\overline{A}_{nj_r})\\
&= 1 - P\Big(\bigcup_{r=1}^{k} A_{nj_r}\Big) - \prod_{r=1}^{k}(1-P(A_{nj_r}))\\
&= 1 - \sum_{r=1}^{k} P(A_{nj_r}) + \sum_{1\leqslant r<s\leqslant k} P(A_{nj_r}A_{nj_s})\\
&\quad - \sum_{1\leqslant r<s<t\leqslant k} P(A_{nj_r}A_{nj_s}A_{nj_t}) + \cdots\\
&\quad - 1 + \sum_{r=1}^{k} P(A_{nj_r}) - \sum_{1\leqslant r<s\leqslant k} P(A_{nj_r})P(A_{nj_s})\\
&\quad + \sum_{1\leqslant r<s<t\leqslant k} P(A_{nj_r})P(A_{nj_s})P(A_{nj_t}) - \cdots\\
&= \sum_{1\leqslant r<s\leqslant k} \{P(A_{nj_r}A_{nj_s}) - P(A_{nj_r})P(A_{nj_s})\}\\
&\quad - \sum_{1\leqslant r<s<t\leqslant k} \{P(A_{nj_r}A_{nj_s}A_{nj_t})\\
&\qquad - P(A_{nj_r})P(A_{nj_s})P(A_{nj_t})\} + \cdots
\end{aligned}$$

Using the definition of random variables X_{nj} and events A_{nj} and equality $u_n = x\sqrt{n}+n$, where $x < 0$, we obtain that

$$\sum_{j=1}^{n} I(A_{nj}) \leqslant C_0(x)\cdot\sqrt{n},$$

where $I(A_{nj})$ is an indicator function: $I(A_{nj}) = 1$ if $X_{nj} > u_n$ holds and $I(A_{nj}) = 0$ otherwise. In other words, the number of exceedances X_{nj} over u_n is at most $O(\sqrt{n})$. The statement of Lemma 4 now follows from Lemma 3. □

Lemma 5. *Let $x < 0$ and $u_n = x\sqrt{n} + n$. Then there exists a real constant $C_2(x)$, such that for positive integers k and l, where $k + l \leqslant n$, and all $1 \leqslant j_1 < j_2 < \cdots < j_k < j_{k+1} < \cdots < j_{k+l} \leqslant n$, the following inequality holds*:

$$\bigg|P\Big(\bigcap_{r=1}^{k+l} \overline{A}_{nj_r}\Big) - P\Big(\bigcap_{r=1}^{k} \overline{A}_{nj_r}\Big)\cdot P\Big(\bigcap_{r=k+1}^{k+l} \overline{A}_{nj_r}\Big)\bigg| \leqslant \frac{C_2(x)}{\sqrt{n}}$$

i.e. Leadbetter's condition $D(u_n)$ is satisfied.

Proof. Lemma 5 is a consequence of Lemma 4 and the following inequality:

$$\begin{aligned}
&\bigg|P\Big(\bigcap_{r=1}^{k+l} \overline{A}_{nj_r}\Big) - P\Big(\bigcap_{r=1}^{k} \overline{A}_{nj_r}\Big)\cdot P\Big(\bigcap_{r=k+1}^{k+l} \overline{A}_{nj_r}\Big)\bigg|\\
&\leqslant \bigg|P\Big(\bigcap_{r=1}^{k+l} \overline{A}_{nj_r}\Big) - \prod_{r=1}^{k+l} P(\overline{A}_{nj_r})\bigg|\\
&\quad + \prod_{r=1}^{k} P(\overline{A}_{nj_r})\cdot\bigg|\prod_{r=k+1}^{k+l} P(\overline{A}_{nj_r}) - P\Big(\bigcap_{r=k+1}^{k+l} \overline{A}_{nj_r}\Big)\bigg|\\
&\quad + \bigg|\prod_{r=1}^{k} P(\overline{A}_{nj_r}) - P\Big(\bigcap_{r=1}^{k} \overline{A}_{nj_r}\Big)\bigg|\cdot P\Big(\bigcap_{r=k+1}^{k+l} \overline{A}_{nj_r}\Big).
\end{aligned}$$

□

Lemma 6. *If $u_n = x\sqrt{n} + n$, $x < 0$, then the following $D'(u_n)$ condition is satisfied*:

$$\lim_{k\to\infty}\limsup_{n\to\infty} n\cdot\sum_{j=2}^{[n/k]} P(A_{n1}A_{nj}) = 0.$$

Proof. Using Lemma 3 we obtain that for every positive integer k

$$n\cdot\sum_{j=2}^{[n/k]} P(A_{n1}A_{nj}) \sim n\cdot\frac{2(-x)^3}{3n^{3/2}} + n\cdot\frac{n}{k}\cdot\frac{x^4}{n^2}$$

$$\sim -\frac{2x^3}{3\sqrt{n}} + \frac{x^4}{k}, \quad n \to \infty,$$

and the condition $D'(u_n)$ follows immediately. □

Proof of Theorem 1. The statement of Theorem 1 follows from Lemma 2, Lemma 5, Lemma 6 and [4] Theorem 3.5.2. □

3. On extreme value distributions. Let us first quote the definition of the domains of attraction of extreme value distributions.

Definition 1. A distribution function F belongs to the domain of attraction of a non-degenerate distribution function G if there exist real constants $a_n > 0$ and b_n, $n \in N$, such that

$$F^n(a_n x + b_n) \to G(x),$$

weakly as $n \to \infty$.

Remark 1. A classical result of Gnedenko [1] states that only three types of distribution functions have non-empty domains of attraction. See [2] for details. The following Fréchet, Weibull and Gumbel distribution functions determine these three types:

$$\Phi_\alpha(x) = \begin{cases} 0, & \text{if } x < 0, \\ \exp(-x^{-\alpha}), & \text{if } x \geqslant 0, \end{cases}$$

$$\Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha), & \text{if } x < 0, \\ 1, & \text{if } x \geqslant 0, \end{cases}$$

$$\Lambda(x) = \exp(-e^{-x}), \qquad -\infty < x < +\infty;$$

where $\alpha > 0$. We refer to Φ_α, Ψ_α and Λ as the extreme value distributions.

Remark 2. Let F_n be the common distribution function of random variables X_{nj} and X^*_{nj}, $j \in \{1, 2, \ldots, n\}$ that were introduced in Sections 1 and 2. The function F_n has a jump at the right end point $x_n := \sup\{t : F_n(t) < 1\} = n - 1$. Consequently, no one of distribution functions F_1, F_2, $F_3, \ldots$ belongs to the domains of attraction of extreme value distributions.

Definition 2. Let $X_{n1}, X_{n2}, \ldots, X_{nk_n}$, $n = 1, 2, \ldots$ be a double array of random variables such that the following conditions are satisfied:

(a) For any n random variables $X_{n1}, X_{n2}, \ldots, X_{nk_n}$ are independent with the common distribution function F_n;

(b) $\lim\limits_{n\to\infty} k_n = +\infty$.

The sequence (F_n) belongs to the domain of attraction of a non-degenerate distribution function G if there exist real constants $a_n > 0$ and b_n, $n \in \mathbf{N}$, such that for every $x \in \mathbf{R}$

$$F_n^{k_n}(a_n x + b_n) \to G(x),$$

weakly as $n \to \infty$. In that case we shall use notation $(F_n) \in \widetilde{D}(G)$.

Remark 3. The sequence (F_n), introduced in Section 2, belongs to the domain of attraction of the Weibull distribution $\Psi_2(x)$. Example of a sequence of distribution functions (F_n) that belongs to the domain of attraction of $\Lambda(x)$ (although no one of distribution functions F_n belongs to the domains of attraction of EV distributions) is given in [5]. We shall use notation F^{-1} for the left continuous inverse of a nondecreasing function F.

Theorem 2. *Let (F_n) be a sequence of distribution functions from Definition* 2. *Suppose that $x_0 := \sup\{t : F_n(t) < 1\}$ does not depend on n. If the following conditions are satisfied*

(a) $x_0 = +\infty$;

(b) $a_n := \left(\dfrac{1}{1-F_n}\right)^{-1}(k_n) \to +\infty$ *as* $n \to \infty$;

(c) $\lim\limits_{n\to\infty} \dfrac{1-F_n(a_n x)}{1-F_n(a_n)} = x^{-\alpha}$ *for any* $x > 0$;

(d) $\lim\limits_{n\to\infty} F_n^{k_n}(0) = 0$;

then $(F_n) \in \widetilde{D}(\Phi_\alpha)$.

Proof. Let the conditions of the theorem are satisfied. Then

$$1 - F_n(a_n) \sim \frac{1}{k_n}, \quad n \to \infty,$$

and consequently we get that for every $x > 0$,

$$k_n(1 - F_n(a_n x)) \sim \frac{1-F_n(a_n x)}{1-F_n(a_n)} \to x^{-\alpha}, \quad n \to \infty.$$

Now we have that for every $x > 0$,

$$F_n^{k_n}(a_n x) \to \exp(-x^{-\alpha}), \quad n \to \infty.$$

If $x < 0$, then $F_n^{k_n}(a_n x) \leqslant F_n^{k_n}(0) \to 0$, as $n \to \infty$. Hence, we proved that for any real x, $F_n^{k_n}(a_n x) \to \Phi_\alpha(x)$, i.e. $(F_n) \in \widetilde{D}(\Phi_\alpha)$. □

Theorem 3. *Let (F_n) be a sequence of distribution functions from Definition* 2 *and*

$$x_n = \sup\{t : F_n(t) < 1\},$$

$$a_n = \left(\frac{1}{1-F_n}\right)^{-1}(k_n).$$

Suppose that the following conditions are satisfied:

(a) $x_n < +\infty$ *for any positive integer* n;

(b) $\lim\limits_{n\to\infty}(x_n - a_n) = 0$;

(c) *There exists* $\alpha > 0$, *such that for every* $t > 0$ *the following equality holds*:

$$\lim_{n\to\infty} \frac{1 - F_n\{x_n - (x_n - a_n)t\}}{1 - F_n(a_n)} = t^\alpha.$$

Then $(F_n) \in \widetilde{D}(\Psi_\alpha)$ *and for every real* t,

$$\lim_{n\to\infty} F_n^{k_n}\{x_n + (x_n - a_n)t\} = \Psi_\alpha(t).$$

Proof. Let us denote

$$F_n^*(x) = \begin{cases} 0, & \text{if } x \leqslant 0, \\ F_n\Big(x_n - \dfrac{1}{x}\Big), & \text{if } x > 0; \end{cases}$$

$$a_n^* = \left(\frac{1}{1 - F_n^*}\right)^{-1}(k_n).$$

Since

$$a_n = \left(\frac{1}{1 - F_n}\right)^{-1}(k_n) = \inf\left\{s : \frac{1}{1 - F_n(s)} \geqslant k_n\right\},$$

and $x_n - a_n \to 0$ as $n \to \infty$, we obtain that

$$\begin{aligned} a_n^* &= \inf\left\{x : \frac{1}{1 - F_n^*(x)} \geqslant k_n\right\} \\ &= \inf\left\{x : \frac{1}{1 - F_n^*\,(x_n - (1/x))} \geqslant k_n\right\} \\ &= \inf\left\{\frac{1}{x_n - s} : \frac{1}{1 - F_n(s)} \geqslant k_n\right\} \\ &= \frac{1}{x_n - a_n} \to \infty, \quad n \to \infty, \end{aligned}$$

and consequently

$$1 - F_n^*(a_n^*) \sim \frac{1}{k_n}, \quad \text{as } n \to \infty.$$

Now, for any $x > 0$ we get

$$\begin{aligned} k_n \cdot \{1 - F_n^*(a_n^* x)\} &\sim \frac{1 - F_n^*(a_n^* x)}{1 - F_n^*(a_n^*)} \\ &= \frac{1 - F_n\,(x_n - (1/a_n^* x))}{1 - F_n\,(x_n - (1/a_n^*))} \\ &= \frac{1 - F_n\,(x_n - (x_n - a_n)(1/x))}{1 - F_n(a_n)} \\ &\to \left(\frac{1}{x}\right)^\alpha = x^{-\alpha}, \quad n \to \infty. \end{aligned}$$

Since all conditions of Theorem 2 are satisfied, we get for all $x > 0$,

$$\lim_{n\to\infty} \{F_n^*(a_n^* x)\}^{k_n} = \exp(-x^{-\alpha}).$$

Let $t < 0$ and $x = -(1/t) > 0$.
In this case we obtain the following relations:

$$\begin{aligned} &F_n^{k_n}\{x_n + (x_n - a_n)t\} \\ &\quad = F_n^{k_n}\left\{x_n - (x_n - a_n)\frac{1}{x}\right\} \\ &\quad = F_n^{k_n}\left(x_n - \frac{1}{a_n^* x}\right) = \{F_n^*(a_n^* x)\}^{k_n} \\ &\quad \to \exp(-x^{-\alpha}) = \exp\left\{-\left(\frac{1}{x}\right)^\alpha\right\} \\ &\quad = \exp\{-(-t)^\alpha\}. \end{aligned}$$

For $t > 0$ we get $F_n^{k_n}\{x_n + (x_n - a_n)t\} = 1$. □

References

[1] Gnedenko, B. V.: Sur la distribution limite du terme maximum d'une série aléatoire. Ann. Math., **44**, 423–453 (1943).

[2] de Haan, L.: On Regular Variation and its Application to the Weak Convergence of Sample Extremes. Mathematical Centre Tracts 32, Mathematisch Centrum, Amsterdam (1970).

[3] Kolchin, V. F.: Random Functions. Nauka, Moscow, (1984). (in Russian).

[4] Leadbetter, M. R., Lindgren, G., and Rootzèn, H.: Extremes and Related Properties of Random Sequences and Processes. Springer-Verlag, New York-Heidelberg-Berlin (1983).

[5] Mladenović, P: Limit theorems for the maximum terms of a sequence of random variables with marginal geometric distributions. Extremes, **2:4**, 405–419 (1999).