CrossMark

J. Appl. Probab. **59,** 825–848 (2022)
*doi:*10.1017/jpr.2021.94

GENERALIZED LIMIT THEOREMS FOR *U*-MAX STATISTICS

YAKOV NIKITIN,* *Saint-Petersburg State University*
EKATERINA SIMAROVA,* ** *Saint-Petersburg State University and Leonhard Euler International Mathematical Institute*

Abstract

U-max statistics were introduced by Lao and Mayer in 2008. Such statistics are natural in stochastic geometry. Examples are the maximal perimeters and areas of polygons and polyhedra formed by random points on a circle, ellipse, etc. The main method to study limit theorems for *U*-max statistics is via a Poisson approximation. In this paper we consider a general class of kernels defined on a circle, and we prove a universal limit theorem with the Weibull distribution as a limit. Its parameters depend on the degree of the kernel, the structure of its points of maximum, and the Hessians of the kernel at these points. Almost all limit theorems known so far may be obtained as simple special cases of our general theorem. We also consider several new examples. Moreover, we consider not only the uniform distribution of points but also almost arbitrary distribution on a circle satisfying mild additional conditions.

Keywords: Weibull distribution; Poisson approximation; *U*-max statistics; random perimeter; random area

2020 Mathematics Subject Classification: Primary 60D05
Secondary 60F05; 60G70

1. Introduction

U-statistics were introduced in probability by Halmos [3] and Hoeffding [4] in the mid-1940s as a functional generalization of sample mean. Let $\xi_1, \xi_2, \dots$ be a sequence of independent identically distributed random elements taking values in a measurable space $(\mathfrak{X}, \mathfrak{A})$. We define a real-valued symmetric Borel function $f(x_1, \dots, x_m)$ on the space $\mathfrak{X}^m$, which we call a kernel of degree m.

U-statistics are defined as follows:

$$U_n = \binom{n}{m}^{-1} \sum_J f(\xi_{i_1}, \dots, \xi_{i_m}), \tag{1.1}$$

where $n \geq m$ and the set

$$J = \{(i_1, \dots, i_m) \colon 1 \leq i_1 < \dots < i_m \leq n\}.$$

Over the past few decades, *U*-statistics have been studied in detail in many publications, and the state of the art is presented in the monographs [7] and [11].

Received 7 November 2020; revision received 20 October 2021.
* Postal address: Department of Mathematics and Mechanics, Universitetsky pr. 28, Stary Peterhof 198504, Russia.
** Postal address: 14th Line 29B, Vasilyevsky Island, St. Petersburg, 199178, Russia. Email address: katerina.1.14@mail.ru

© The Author(s), 2022. Published by Cambridge University Press on behalf of Applied Probability Trust.

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

In 2008 Lao and Meyer [9, 10, 14] independently considered the so-called U-max statistics, which are obtained from (1.1) by removing the normalizing factor and replacing the sum with the maximum over the set J:

$$M_n = \max_J f(\xi_{i_1}, \dots, \xi_{i_m}). \tag{1.2}$$

The U-min statistics M'_n are defined in a similar way by replacing max with min.

Lao and Mayer mainly studied the limit behavior of maximal and minimal distances, areas and perimeters of sets formed by random points on a circle or a sphere. They used the Poisson approximation from the monograph [2] and the paper [16], and proved a number of theorems on convergence to a Weibull distribution.

Lao and Mayer studied the kernels of low degrees, for example the area and perimeter of random triangles (see [10]). Koroleva and Nikitin considered U-max statistics of a more complicated nature (see [6]). In particular, they considered the maximal perimeter among all perimeters of convex m-gons, where random vertices are chosen from n independent points uniformly distributed on a circle. This was generalized in another direction in the papers [17] and [18], where a generalized perimeter of a random convex polygon was considered.

These papers considered the uniform distribution of points on the unit circle. More general distributions were used by Lao and Mayer for some particular two-dimensional kernels, namely for the distances between points and the scalar product of two position vectors in [9], [10], and [14]. More complicated kernels were studied in [15], namely the areas and perimeters of inscribed polygons with more general conditions on the distribution of vertices. A common feature of all papers was that they investigated specific particular cases of U-max statistics and U-min statistics.

This paper is devoted to a significant generalization of known limit theorems for U-max and U-min statistics. We consider an almost arbitrary distribution of points on a circle, as well as a wide and general class of smooth kernels with a natural structure of the set of extreme points. In this formulation, the limit behavior is determined by the degree of the kernel f, by the Hessian of the kernel at the maximal points, and also by the distribution of the points on the circle. The general formulas are used for kernels of a special type with convexity properties. Most of the previously known results can be deduced from Theorem 3.1 but we also provide some new examples.

Our paper consists of several parts. First, in Section 2, we introduce the basic notation and restrictions, and then in Section 3 we formulate the general limit relations we have obtained. Further, in Sections 4 and 5 we apply our results to some specific classes of kernels. For these classes, more explicit and relatively simple limit theorems for U-max statistics of geometric nature are obtained. Some interesting examples are also given. A detailed proof of the main result of Section 3 is rather painstaking and is of considerable length. Therefore we place it at the end of the paper; it occupies Sections 6 and 7.

2. Preliminaries

In this section we introduce necessary conditions and definitions. We consider U-max statistics with a fixed kernel f depending on m points $V_1, \dots, V_m$ lying on the unit circle S^1 with center O at the origin, that is,

$$f\colon (S^1)^m \to \mathbb{R}.$$

Further, we always assume that $m \geq 2$.

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

Let β_i denote the angle between the vectors OV_1 and OV_{i+1} (taken counterclockwise). We call such angles central. In this way,

$$\beta_i = \angle V_1 O V_{i+1}. \tag{2.1}$$

Sometimes for the sake of brevity we will use the notation

$$\beta = (\beta_1, \dots, \beta_{m-1}) \in [0, 2\pi)^{m-1}. \tag{2.2}$$

All angles that appear in this paper are considered modulo 2π. All algebraic operations involving several angles are also considered modulo 2π, unless otherwise stated.

Now we give some conditions that will be used below.

A. Conditions on the kernel f.

A1. The function f is invariant with respect to rotations. Equivalently, this means that the function f can be written in the form

$$f(V_1, \dots, V_m) = h(\beta_1, \dots, \beta_{m-1}) = h(\beta), \tag{2.3}$$

where β_i are central angles, and $h\colon [0, 2\pi)^{m-1} \to \mathbb{R}$ is a function.

A2. The function f cannot be changed after any permutation of the points $V_1, \dots, V_m$. Therefore the function h is also a symmetric function of its arguments.

A3. The function h is continuous and can be continuously extended to a function $h\colon [0, 2\pi]^{m-1} \to \mathbb{R}$.

A4. The function h reaches its maximal value M and this maximum is realized only at a finite number of points $W_1, \dots, W_k \in [0, 2\pi]^{m-1}$. It is assumed that none of these points lie on the boundary of the domain of definition of the function h. In other words, $W_i^j \in (0, 2\pi)$ for all $i \in \{1, \dots, k\}$, $j \in \{1, \dots, m-1\}$, where W_i^j is the jth component of the point W_i.

Condition A4 together with condition A2 allows us to make the following conclusion about the structure of the maxima of the function f: there is only a finite number of points (up to rotations) where the maximal value of the function f is attained. Moreover, none of these points have matching components.

A5. There exists $\delta > 0$ such that the function h is three-times continuously differentiable in the δ-neighborhood of any maximum point W_i, $i \in \{1, \dots, k\}$.

A6. Consider the Hessian matrix G_i of the form

$$G_i = \begin{pmatrix} \dfrac{\partial^2 h(W_i)}{\partial^2 x_1} & \dfrac{\partial^2 h(W_i)}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 h(W_i)}{\partial x_1 \partial x_{m-1}} \\ \dfrac{\partial^2 h(W_i)}{\partial x_1 \partial x_2} & \dfrac{\partial^2 h(W_i)}{\partial^2 x_2} & \cdots & \dfrac{\partial^2 h(W_i)}{\partial x_2 \partial x_{m-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 h(W_i)}{\partial x_{m-1} \partial x_1} & \dfrac{\partial^2 h(W_i)}{\partial x_{m-1} \partial x_2} & \cdots & \dfrac{\partial^2 h(W_i)}{\partial^2 x_{m-1}} \end{pmatrix}.$$

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

We require that the condition

$$\det(G_i) \neq 0$$

holds for all $i \in \{1, \dots, k\}$.

B. Conditions on the distribution of points.

B1. The random points $V_1, \dots, V_n$ are independently distributed on the unit circle S^1 with the same probability density function $p(x)$.

B2. The density function p is continuous (therefore it can be considered as a non-negative continuous 2π-periodic function $p\colon \mathbb{R} \to \mathbb{R}_+$ such that $\int_0^{2\pi} p(x)\mathrm{d}x = 1$).

B3. There exists at least one maximal point of the kernel (which we denote by W_*) such that

$$\int_0^{2\pi} \left[p(x) \prod_{l=1}^{m-1} p\big(x + W_*^l\big) \right] \mathrm{d}x \neq 0.$$

Similar conditions on $p(x)$ arose in [15].

Remark 2.1. The conditions imposed on the density p are not too restrictive. For example, continuous densities separated from zero or continuous densities taking the value 0 only on a set of measure less than $2\pi/m$ are suitable for these conditions. A useful example of non-uniform distribution is the von Mises distribution (see e.g. [13]).

3. Main results

Now we state the main result of this paper.

Theorem 3.1. (General theorem.) *Suppose that a kernel f and points $V_1, \dots, V_n$ satisfy all conditions* A *and* B. *Let M_n be the U-*max *statistic constructed by a kernel f, that is, $M_n = \max_{1 \le i_1 < \dots < i_m \le n} f(V_{i_1}, \dots, V_{i_m})$.*

Then for every $t > 0$ the following relation holds true:

$$\lim_{n\to\infty} \mathbb{P}\big\{ n^{2m/(m-1)}(M - M_n) \le t \big\} = 1 - \exp\left(-\frac{t^{(m-1)/2} K_m}{m!} \right),$$

where

$$K_m = \frac{(2\pi)^{(m-1)/2}}{\Gamma\big(\frac{m+1}{2}\big)} \sum_{i=1}^{k} \left(\frac{1}{\sqrt{\det(-G_i)}} \int_0^{2\pi} p(x) \prod_{l=1}^{m-1} p\big(x + W_i^l\big)\, \mathrm{d}x \right)$$

and M is from condition A4. *The rate of convergence is* $\mathrm{O}\big(n^{-1/(m-1)}\big)$.

Theorem 3.1 immediately implies several simple but very useful consequences. As far as we know, these consequences are new. First of all, Theorem 3.1 can be modified slightly for U-min statistics.

Corollary 3.1. *Let us replace the maximum M with the minimum μ and consider the points of minimum in conditions* A4, A5, A6, *and* B3. *Let M_n' denote the U-*min *statistic constructed by a kernel f, i.e. $M_n' = \min_{1 \le i_1 < \dots < i_m \le n} f(V_{i_1}, \dots, V_{i_m})$.*

Then for each $t > 0$ we have

$$\lim_{n\to\infty} \mathbb{P}\big\{ n^{2m/(m-1)}(M_n' - \mu) \le t \big\} = 1 - \exp\left(-\frac{t^{(m-1)/2} K_m}{m!} \right),$$

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

where

$$K_m = \frac{(2\pi)^{(m-1)/2}}{\Gamma\left(\frac{m+1}{2}\right)} \sum_{i=1}^{k} \left(\frac{1}{\sqrt{\det(G_i)}} \int_0^{2\pi} p(x) \prod_{l=1}^{m-1} p\left(x + W_i^l\right) \mathrm{d}x \right).$$

Corollary 3.2. (Particular case.) *If p is the uniform density, then the constant K_m from Theorem 3.1 has the following form:*

$$K_m = \frac{1}{(2\pi)^{(m-1)/2}\Gamma\left(\frac{m+1}{2}\right)} \sum_{i=1}^{k} \frac{1}{\sqrt{\det(-G_i)}}.$$

Remark 3.1. Theorem 3.1 may be applied to a rather wide class of kernels from stochastic geometry but of course not to all of them. In [18] the definition of a generalized perimeter was introduced. It is the sum of the yth degrees of the side lengths of a polygon constructed on m given points on a circle. It is shown that for $y > 1$ and $m > 1 + \pi/(\arccos 1/\sqrt{y})$, the generalized perimeter attains its maximum on the configuration of points in which some of them coincide. This shows that our result cannot be applied to this rather simple variant of the problem. In this case the limit relation is still an open question.

Remark 3.2. Note that condition A3 may be replaced with the following one: the function h is continuous and can be continuously extended to a function $h\colon [0, 2\pi]^{m-1} \to \mathbb{R} \cup \{-\infty\}$. It follows from the fact that the function $\hat{h}(x) = \max(h(x), M - 1)$, where M is from condition A4 and $\hat{f}$ constructed from $\hat{h}$ by (2.3), satisfies Theorem 3.1. The limiting relation does not change after replacing $\hat{f}$ with f.

In the following sections we give limit theorems for U-max statistics generated by kernels of a special form that attain their maximum only at the vertices of a regular polygon. This special case covers a large number of standard geometric characteristics.

4. Limit behavior of U-max statistics for several functions depending on the side lengths of the polygon

Suppose that $V_1, \ldots, V_m$ and $\beta_1, \ldots, \beta_{m-1}$ are as introduced earlier, and also for a moment assume that $0 = \beta_0 \le \beta_1 \le \ldots \le \beta_{m-1} \le \beta_m = 2\pi$. We construct functions f and h via the following formula:

$$f(V_1, \ldots, V_m) = h(\beta_1, \ldots, \beta_{m-1}) = \sum_{i=1}^{m} g(\beta_i - \beta_{i-1}), \tag{4.1}$$

where $g\colon [0, 2\pi] \to \mathbb{R}$ is continuous and three-times continuously differentiable in some neighborhood of the point $2\pi/m$. We also require $g''(2\pi/m) \neq 0$. Further, we define the function h on $[0, 2\pi]^{m-1}$ so that it is symmetric and corresponds to the function f by (2.3). Such a function f satisfies conditions A1 and A3. Condition A2 also holds since the function f depends only on the angles between neighboring vectors $OV_1, \ldots, OV_m$.

Under additional restrictions we get the following statement.

Theorem 4.1. *Suppose that the points $V_1, \ldots, V_n$ are independently distributed on S^1 with a common continuous density $p(x)$ such that*

$$\int_0^{2\pi} \prod_{l=0}^{m-1} p\left(x + \frac{2\pi l}{m}\right) \mathrm{d}x > 0.$$

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

Consider the U-max *statistic* M_n *with a kernel* f *of the form* (4.1). *Suppose that this kernel attains its maximum only at the vertices of a regular* m*-gon.*

Then for any $t > 0$ *the following limit relation holds:*

$$\lim_{n\to\infty} \mathbb{P}\left\{ n^{2m/(m-1)}\left(mg\left(\frac{2\pi}{m}\right) - M_n \right) \le t \right\} = 1 - \exp\left(-\frac{t^{(m-1)/2} K_m}{m} \right), \tag{4.2}$$

where

$$K_m = \frac{(2\pi)^{(m-1)/2}\left[\int_0^{2\pi} \prod_{l=0}^{m-1} p\left(x + \frac{2\pi l}{m}\right) \mathrm{d}x\right]}{\left(-g''\left(\frac{2\pi}{m}\right)\right)^{(m-1)/2} \Gamma\left(\frac{m+1}{2}\right)\sqrt{m}}.$$

Proof. Let us first prove that the function f and density p satisfy conditions A and B from Section 3. It was established above that conditions A1, A2, and A3 are met. The fulfillment of condition A4 follows from the fact that a regular polygon is the only maximal point of the function f. Condition A5 follows from formula (4.1) and the differentiability assumption. Conditions B1 and B2 are obviously satisfied (they are assumed in the statement of the theorem). It remains to check the properties A6 and B3; then we can use Theorem 3.1.

By (2.2), a regular m-gon corresponds to some permutation of the angles in $W^* = (W^1, \dots, W^{m-1})$, where $W^i = 2\pi i/m$. Thus there are $(m-1)!$ maxima of the corresponding function h, and the Hessian determinants of the function h at all these points are the same. It allows us to restrict ourselves to the case $0 \le \beta_1 \le \dots \le \beta_{m-1} \le 2\pi$. Thus the condition that the maximum of the function f is attained only on a regular m-gon means that the point W^* is the maximum of the function h among all points with ordered angles. Together with the condition

$$\int_0^{2\pi} \prod_{l=0}^{m-1} p\left(x + \frac{2\pi l}{m}\right) \mathrm{d}x > 0,$$

it implies the validity of condition B3. Next we obtain an explicit formula for the determinant of the Hessian matrix of the function h at the point W^* and make sure that it is not equal to zero. This fact implies the fulfillment of condition A6, and then an application of Theorem 3.1 finishes the proof of Theorem 4.1.

By simple calculations, we get

$$\frac{\partial^2 h}{\partial x_i \partial x_j}(W^*) = \begin{cases} 0 & \text{if } |i-j| > 1, \\ 2g''\left(\dfrac{2\pi}{m}\right) & \text{if } i = j, \\ -g''\left(\dfrac{2\pi}{m}\right) & \text{if } |i-j| = 1. \end{cases}$$

Therefore the Hessian matrix at the point W^* is

$$G(W^*) = g''\left(\frac{2\pi}{m}\right) \begin{pmatrix} 2 & -1 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 2 \end{pmatrix} = g''\left(\frac{2\pi}{m}\right) B_{m-1}.$$

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

The determinant of the tridiagonal Toeplitz matrix B_{m-1} is well known: it is equal to m. Hence

$$\det(-G(W^*)) = m \cdot \left(-g''\left(\frac{2\pi}{m}\right)\right)^{m-1} \neq 0,$$

and the conditions of Theorem 3.1 are satisfied. Recalling the fact that there are $(m-1)!$ maxima of the function h with the same determinant of the Hessian matrix and the same integral, we obtain the desired limit relation. □

Lemma 4.1. *For strictly concave functions* $g\colon [0, 2\pi] \to \mathbb{R}$, *the maximum of the function* f *defined by* (4.1) *is attained only at the vertices of a regular* m *-gon.*

Proof. We assume that $0 \le \beta_1 \le \ldots \le \beta_{m-1} \le 2\pi$. Let us prove that the point $W^* = (W^1, \ldots, W^{m-1})$, where $W^i = 2\pi i/m$, is the only maximum of the function f among all points with ordered angles. Due to Jensen's inequality we have

$$f(V_1, \ldots, V_m) = \sum_{i=1}^{m} g(\beta_i - \beta_{i-1}) \le mg\left(\frac{\beta_m - \beta_0}{m}\right) = mg\left(\frac{2\pi}{m}\right) = \sum_{i=1}^{m} g\big(W^i - W^{i-1}\big). \tag{4.3}$$

The function g is strictly concave; therefore, if the arguments of the function g are not all equal, inequality in (4.3) is strict. □

By combining Theorem 4.1 and Lemma 4.1, we obtain the following corollary.

Corollary 4.1. *Suppose that a function* $g\colon [0, 2\pi] \to \mathbb{R}$ *is continuous, strictly concave, three-times continuously differentiable in a neighborhood of the point* $2\pi/m$ *and* $g''(2\pi/m) \neq 0$. *Let the function* f *be defined by* (4.1). *Then the* U-max *statistic with a kernel* f *satisfies relation* (4.2) *of Theorem* 4.1.

Corollary 4.2. *In Theorem* 4.1 *we may consider functions* f *for which a regular polygon is the only point of minimum. Then, for the* U-min *statistic* M'_n *generated by a kernel* f, *the following limit relation holds for any* $t > 0$ (*similarly to* (4.2)):

$$\lim_{n\to\infty} \mathbb{P}\left\{n^{2m/(m-1)}\left(M'_n - mg\left(\frac{2\pi}{m}\right)\right) \le t\right\} = 1 - \exp\left(-\frac{t^{(m-1)/2}K_m}{m}\right),$$

where

$$K_m = \frac{(2\pi)^{(m-1)/2}\big[\int_0^{2\pi} \prod_{l=0}^{m-1} p\big(x + \frac{2\pi l}{m}\big)\mathrm{d}x\big]}{\big(g''\big(\frac{2\pi}{m}\big)\big)^{(m-1)/2}\Gamma\big(\frac{m+1}{2}\big)\sqrt{m}}.$$

In particular, this is the case when we consider the strictly convex functions g *in Corollary* 4.1 *instead of strictly concave ones.*

Remark 4.1. If p is the uniform density, then the constant K_m from Theorem 4.1 satisfies

$$K_m = \frac{1}{\big(-2\pi g''\big(\frac{2\pi}{m}\big)\big)^{(m-1)/2}\Gamma\big(\frac{m+1}{2}\big)\sqrt{m}}. \tag{4.4}$$

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

Example 4.1. (*Maximal perimeter of inscribed polygon.*) Let us consider the maximal perimeter of an inscribed convex m-gon with random vertices on a circle with $m \geq 3$. This is the U-max statistic with a kernel f of the form (4.1), where $g(x) = 2\sin(x/2)$. This function is strictly concave, hence the results follow from Theorem 4.1. The limit behavior of the U-max statistic with such a kernel in the case of the uniform distribution of points may be found in [6]; the result coincides with Remark 4.1 with this function g, so the result [6] is a simple special case of our results.

Example 4.2. (*Maximal area of inscribed polygon.*) Another U-max statistic considered in [6] was the area of an inscribed convex m-gon with $m \geq 3$. It is generated by a kernel f of the form (4.1), where $g(x) = \frac{1}{2}\sin x$. The maximum of the function f is attained only on a regular m-gon; see e.g. [20, problem 57a]. Therefore the limit theorem for this U-max statistic follows directly from Theorem 4.1 and Remark 4.1. The results of [6] and more general results of [15] again follow from ours as simple special cases. Similar statements hold for the areas and perimeters of the *described* random polygons considered in [6]; they also follow from Theorem 4.1 with some $g(x)$.

Example 4.3. (*Sum of the distances from the center to the vertices of described polygon.*) Let us now consider the example of a kernel that did not arise earlier in the literature on the limit behavior of U-max statistics. We define a kernel $f\colon (S^1)^m \to \mathbb{R} \cup \{+\infty\}$ with $m \geq 3$ as follows: construct the described convex m-gon with vertices at points $A_1, \ldots, A_m$ such that its sides touch the circle S^1 at points $U_1, \ldots, U_m$. Define the function

$$f(V_1, \ldots, V_m) = \sum_{i=1}^{m} |OA_i|$$

as the sum of distances from the center to the vertices of the described m-gon. It is possible that some vertex A_i goes to infinity; in this case we define $f(V_1, \ldots, V_m) = +\infty$.

The function f can be written in the form (4.1) when $g(x) = (\cos(x/2))^{-1}$, if $0 \leq x < \pi$, and $g(x) = +\infty$ otherwise. The case $g(x) = +\infty$ can occur if some vertex A_i goes to infinity. Note that

$$g''(x) = \frac{1 + \sin^2(x/2)}{4\cos^3(x/2)} > 0 \quad \text{for } x \in [0, \pi),$$

so the function is strictly convex on $[0, \pi)$. By Corollary 4.2, the minimum will be attained at the vertices of a regular m-gon, and for the U-min statistic M'_n generated by a kernel f we have for any $t > 0$

$$\lim_{n\to\infty} \mathbb{P}\left\{ n^{2m/(m-1)}\left(M'_n - \frac{m}{\cos\frac{\pi}{m}} \right) \leq t \right\} = 1 - \exp\left(-\frac{t^{(m-1)/2} K_m}{m} \right),$$

where

$$K_m = \frac{\left(8\pi\cos^3\frac{\pi}{m}\right)^{(m-1)/2} \left[\int_0^{2\pi} \prod_{l=0}^{m-1} p\left(x + \frac{2\pi l}{m}\right) \mathrm{d}x\right]}{\left(1 + \sin^2\frac{\pi}{m}\right)^{(m-1)/2} \Gamma\left(\frac{m+1}{2}\right)\sqrt{m}}.$$

Example 4.4. (*Generalized perimeter of the polygon.*) In [17, 18] a definition of a generalized perimeter was introduced. A generalized perimeter of order y is the sum of the yth degrees of the side lengths of a convex inscribed polygon constructed on m given points. This function may also be written as

$$f(V_1, \ldots, V_m) = \sum_{i=1}^{m} g(\beta_i - \beta_{i-1}),$$

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

where $g(x) = 2^y \sin^y(x/2)$. The function g is convex for negative y and concave for $y \in (0, 1]$, so the limit relation obtained in these cases for U-min and U-max statistics, respectively, is also a special case of the equality (4.4) (see [17]).

For $y \in (1, 2]$ and $m = 3$, the function g is not concave but the function f also attains its maximum only on the vertices of regular triangle, so relation (4.2) with constant (4.4) from Theorem 4.1 also holds true in this case. This was shown in [18].

Example 4.5. (*Further generalization of the perimeter.*) The concept of a generalized perimeter introduced in [17] may be generalized further. Suppose that a kernel f is given by

$$f(V_1, \dots, V_m) = \sum_{i=1}^{m} r(|V_i V_{i+1}|), \tag{4.5}$$

where $r\colon [0, 2] \to \mathbb{R}$ is a function and the points $V_1, \dots, V_m$ are ordered counterclockwise, and we also assume that the kernel is symmetric. By $|V_i V_{i+1}|$ we denote the length of the side of the polygon. The generalized perimeter introduced in [17] and [18] corresponds to the function $r(x) = x^y$.

If the function r is continuous, strictly concave, increasing, three-times continuously differentiable in some neighborhood of the point $2\sin(\pi/m)$ and $r''(2\sin(\pi/m)) \neq 0$, then the function $g(x) = r(2\sin(x/2))$ is strictly concave. By Corollary 4.1, we can write the limit relation for U-max statistics from Theorem 4.1. It is also possible to replace the conditions of strict concavity and increasing with the conditions of concavity and strict increasing.

Similarly, if r is a continuous strictly convex decreasing function which is three-times continuously differentiable in some neighborhood of the point $2\sin(\pi/m)$, and $r''(2\sin(\pi/m)) \neq 0$, then the function $g(x) = r(2\sin(x/2))$ is strictly convex. By Corollary 4.2, we can write the limit relation for U-min statistics. The condition of strict convexity and decreasing may be replaced by the condition of convexity and strict decreasing.

In particular, let us use the function $r(x) = \mathrm{e}^{-ax} x^b (\ln(x/2))^c$. Such functions were considered by Alexander and Stolarsky in [1]. Let $\tau(a, b, c)$ denote the function that is equal to 1 if the function r is strictly concave and increasing, and equal to -1 if the function r is strictly convex and decreasing. Then the following equality from [1] is true:

$$\tau(a, b, c) = \begin{cases} (-1)^c & \text{if } a \geq 0,\ b \leq 0,\ c \in \mathbb{N}, \\ -1 & \text{if } a \geq 0,\ b \leq 0,\ c = 0,\ a^2 + b^2 \neq 0, \\ 1 & \text{if } a = 0,\ 0 < b \leq 1,\ c = 0. \end{cases}$$

We may study the limit behavior of U-max (resp. U-min) statistics M_n (resp. M_n') with a kernel f constructed by (4.5) in the case when $\tau = 1$ (resp. $\tau = -1$). It can be simply done by using $g(x) = r(2\sin(x/2))$ in Theorem 4.1.

5. Limit behavior of U-max statistics for several functions depending on the side lengths and diagonals of a polygon

In this section the arguments are very similar to those in Section 4. We define a kernel f and the corresponding function h by the set of angles $(\beta_1, \dots \beta_{m-1})$ introduced in (2.2). The connection between f and h is given by (2.3). We again define the function h using a

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

continuous function $g\colon [0, 2\pi] \to \mathbb{R}$, but in a different way. The restrictions on the function g will be different, and the function h itself is defined via the following analogue of (4.1):

$$f(V_1, \dots, V_m) = h(\beta_1, \dots, \beta_{m-1}) = \sum_{0 \le i < j \le m-1} g(|\beta_j - \beta_i|), \quad \text{where } \beta_0 = 0. \tag{5.1}$$

In other words, in this section we consider *functions depending on the angles between any pairs of points V_i and V_j, but not only on the angles between adjacent points,* as was done in the previous section.

We state an analogue of Theorem 4.1 for a kernel f of the form (5.1).

Theorem 5.1. *Suppose that the points $V_1, \dots, V_n$ are independently distributed on S^1 with a common continuous density function $p(x)$ such that*

$$\int_0^{2\pi} \prod_{l=0}^{m-1} p\left(x + \frac{2\pi l}{m}\right) \mathrm{d}x > 0.$$

*Consider a continuous function $g\colon [0, 2\pi] \to \mathbb{R}$ which is three-times continuously differentiable in the neighborhoods of points $2\pi s/m$ for all $s \in \{1, \dots, m-1\}$ and such that $g(x) = g(2\pi - x)$. We construct a function f by (5.1) and suppose that it attains its maximum only at the vertices of a regular m-gon. Let M_n be the U-*max *statistic with a kernel f. Consider the symmetric matrix $G = (g_{i,j})_{i,j=1}^{m-1}$, where*

$$g_{i,j} = \begin{cases} -g''\left(\dfrac{2\pi|i-j|}{m}\right) & \text{if } i \ne j, \\ \displaystyle\sum_{s=1}^{m-1} g''\left(\frac{2\pi s}{m}\right) & \text{if } i = j. \end{cases} \tag{5.2}$$

If $\det G \ne 0$, *then, for any $t > 0$, the following limit relation holds:*

$$\lim_{n\to\infty} \mathbb{P}\left\{ n^{2m/(m-1)} \left(\frac{1}{2} \sum_{s=1}^{m-1} m g\left(\frac{2\pi s}{m}\right) - M_n \right) \le t \right\} = 1 - \exp\left(-\frac{t^{(m-1)/2} K_m}{m} \right), \tag{5.3}$$

where

$$K_m = \frac{(2\pi)^{(m-1)/2} \left[\int_0^{2\pi} \prod_{l=0}^{m-1} p\left(x + \frac{2\pi l}{m}\right) \mathrm{d}x \right]}{\sqrt{\det(-G)}\, \Gamma\left(\frac{m+1}{2}\right)}.$$

The proof of this theorem is similar to the proof of Theorem 4.1 and therefore omitted.

Remark 5.1. If the values of second derivatives at the points $2\pi s/m$ are negative for all $s \in \{1, \dots, m-1\}$, then

$$g_{i,i} > \sum_{j=1 i \ne j}^{m-1} |g_{i,j}|.$$

Matrices with such a property are called diagonally dominant, and according to [5, Chapter 6, § 1, p. 392, Th. 6.1.10] the determinants of such matrices are non-zero.

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

Now we describe the extreme points of the function f for concave g having some symmetry property. This is an analogue of Lemma 4.1.

Lemma 5.1. *Suppose that a function $g\colon [0, 2\pi] \to \mathbb{R}$ is a continuous and strictly concave function such that $g(x) = g(2\pi - x)$. Then the function f defined by (5.1) attains its maximum only at the vertices of a regular m-gon, and its maximal value is equal to*

$$\frac{1}{2}\sum_{s=1}^{m-1} mg\left(\frac{2\pi s}{m}\right).$$

Proof. Without loss of generality, we assume that the vertices $V_1, \dots, V_m$ are ordered counterclockwise. Let $P(k, V_1, \dots, V_m)$ denote the sum $\sum_{i=1}^{m} g(\beta_{i+k} - \beta_i)$ assuming that $\beta_0 = 0$, $\beta_{m+s} = 2\pi + \beta_s$ for $s \in \{0, \dots, m-1\}$. Then

$$f(V_1, \dots, V_m) = \frac{1}{2}\sum_{i=1}^{m-1} P(i, V_1, \dots, V_m).$$

We prove that the maximum of $P(k, V_1, \dots, V_m)$ is attained only on a regular m-gon.

Due to strict concavity,

$$P(k, V_1, \dots, V_m) = \sum_{i=1}^{m} g(\beta_{i+k} - \beta_i) \le mg\left(\frac{2\pi k}{m}\right),$$

and equality is achieved only when $\beta_{i+k} - \beta_i = 2\pi k/m$ for all i and k. Therefore the maximal value is attained only at the vertices of a regular polygon. □

Combining the results of Theorem 5.1 and Lemma 5.1, we obtain the following corollary.

Corollary 5.1. *If a function $g\colon [0, 2\pi] \to \mathbb{R}$ is continuous, strictly concave, three-times continuously differentiable in the neighborhoods of points $2\pi s/m$ for all $s \in \{1, \dots, m-1\}$, and also has the property $g(x) = g(2\pi - x)$, and if the probability density function p is continuous and*

$$\int_0^{2\pi} \prod_{l=0}^{m-1} p\left(x + \frac{2\pi l}{m}\right) \mathrm{d}x > 0,$$

*then for the U-*max *statistic M_n with a kernel f of the form (5.1), relation (5.3) of Theorem 5.1 holds.*

Similarly to Section 4, this statement can be reformulated for U-min statistics and strictly convex functions g.

Remark 5.2. One of the main differences between Sections 5 and 4 is the condition $g(x) = g(2\pi - x)$, which was not involved in Section 4. It gives the equality of elements on the main diagonal of the matrix G introduced in Theorem 5.1. This condition also implies that $g(|\beta_i - \beta_j|) = r(|V_iV_j|)$, where $|V_iV_j|$ is the length of the segment V_iV_j.

Thus, by analogy with Example 4.5, we may say that all the functions f satisfying (5.1) are the generalized sums of pairwise distances between points. This shows that the functions g from Examples 4.1 and 4.3 cannot be used to construct the functions f in (5.1). This property holds for the function g from Examples 4.1, 4.4, and 4.5. Such functions g are also differentiable on $(0, 2\pi)$, and Remark 5.1 holds for them. Therefore the functions f constructed in these examples by (5.1) satisfy Theorem 5.1.

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

Example 5.1. As already mentioned, the function $g(x)$ introduced in Example 4.5 may be used in this case as well.

Example 5.2. (*Sum of pairwise distances between vertices.*) Consider the simple case where $g(x)$ is defined in Example 4.1, and the resulting function $f(V_1, \dots, V_m)$ is the sum of the pairwise distances between points $V_1, \dots, V_m$ with $m \geq 3$. According to [19], the maximum is attained only at the vertices of a regular polygon and is equal to $m \cot(\pi/2m)$. The limit relation (5.3) holds but we did not manage to calculate the exact constant K_m in it for an arbitrary m. The elements of the matrix G involved in Theorem 5.1 have the following form:

$$g_{i,j} = \begin{cases} \frac{1}{2} \sin \frac{|i-j|\pi}{m} & \text{if } i \neq j, \\ -\frac{1}{2} \cot \frac{\pi}{2m} & \text{if } i = j. \end{cases}$$

Example 5.3. (*Pairwise sum of inverse distances between vertices.*) Let $g(x) = (2\sin(x/2))^{-1}$ and construct a kernel f using formula (5.1) with $m \geq 3$. It is easy to see that the constructed kernel f is equal to

$$f(V_1, \dots, V_m) = \sum_{j>i} \frac{1}{|V_i V_j|},$$

that is, the sum of the inverse distances between the vertices $V_1, \dots, V_m$, where $m \geq 3$. This example was considered by Toth [19].

The minimal value of a kernel f is attained only at the vertices of a regular polygon and is equal to

$$\frac{m}{4} \sum_{k=1}^{m-1} \frac{1}{\sin\left(\frac{\pi k}{m}\right)}$$

(see [19]). The second derivative of the function g is equal to

$$g''(x) = \frac{2 - \sin^2 \frac{x}{2}}{8 \sin^3 \frac{x}{2}},$$

so the elements of the matrix G defined by (5.2) have the following form:

$$g_{i,j} = \begin{cases} -\dfrac{2 - \sin^2\left(\frac{\pi|i-j|}{m}\right)}{8 \sin^3\left(\frac{\pi|i-j|}{m}\right)} & \text{if } i \neq j, \\ \displaystyle\sum_{s=1}^{m-1} \frac{2 - \sin^2\left(\frac{\pi s}{m}\right)}{8 \sin^3\left(\frac{\pi s}{m}\right)} & \text{if } i = j. \end{cases}$$

6. Proof of Theorem 3.1: part 1

We return to the proof of our general Theorem 3.1. It can be divided into two parts. The first part takes the form of the following statement.

Theorem 6.1. *Suppose that a kernel f and points $V_1, \dots, V_n$ satisfy conditions* A *and* B *from Section 2. Let M_n be the U-*max *statistic constructed via a kernel f, that is,*

$$M_n = \max_{1 \leq i_1 < \dots < i_m \leq n} f(V_{i_1}, \dots, V_{i_m}).$$

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

Then the following two statements are true.

(i) *There exist constants $C > 0$ and $D > 0$ such that, for any number ε, $0 < \varepsilon < C$, if $f(V_1, \dots, V_m) \geq M - \varepsilon$, then $\min_{1 \leq i \leq k} \|W_i - \beta\| \leq D\sqrt{\varepsilon}$, where β is defined by (2.2) and (2.1), and W_i is defined by condition* A4.

(ii) *The following relation holds true:*

$$\lim_{\varepsilon \to 0+} \varepsilon^{-(m-1)/2} \mathbb{P}\{f(V_1, \dots, V_m) \geq M - \varepsilon\} = K_m,$$

where

$$K_m = \frac{(2\pi)^{(m-1)/2}}{\Gamma\left(\frac{m+1}{2}\right)} \sum_{i=1}^{k} \left(\frac{1}{\sqrt{\det(-G_i)}} \int_0^{2\pi} \left(p(x) \prod_{l=1}^{m-1} p\left(x + W_i^l\right) \right) \mathrm{d}x \right).$$

Proof. It is clear that

$$\mathbb{P}\{f(V_1, \dots, V_m) > z\} = \mathbb{P}\{h(\beta_1, \dots, \beta_{m-1}) > z\},$$

where β_i are random angles defined by (2.1). Further, we deal with the function h only.

Let us define, for every $\varepsilon > 0$, the number

$$\hat{S}(\varepsilon) = \min\{s \geq 0 \mid \forall x \in [0, 2\pi]^{m-1} \colon M - f(x) \leq \varepsilon \Rightarrow \exists i \colon \|x - W_i\| \leq s\}. \tag{6.1}$$

In other words $\hat{S}(\varepsilon)$ is the minimal radius of balls with centers in W_i, $i = 1, \dots, k$, such that if the value of the function differs from the maximum value by less than ε, then the argument of the function must lie in one of the balls. For any $\varepsilon > 0$, this minimal radius obviously exists. Also, let us define for any $\varepsilon > 0$

$$S(\varepsilon) = \max\left(\hat{S}(\varepsilon), \varepsilon^{1/3}\right). \tag{6.2}$$

It is easy to show that

$$\lim_{\varepsilon \to 0+} S(\varepsilon) = 0. \tag{6.3}$$

Indeed, (6.3) is equivalent to the limit relation $\lim_{\varepsilon \to 0+} \hat{S}(\varepsilon) = 0$. The function $\hat{S}(x)$ is non-decreasing and non-negative, so the limit $\lim_{\varepsilon \to 0+} \hat{S}(\varepsilon)$ exists and is non-negative. We denote it by a and suppose that $a > 0$ (otherwise (6.3) is proved). Then for any $\varepsilon > 0$ there exists $x_\varepsilon \in [0, 2\pi]^{m-1}$ such that $\min_{1 \leq i \leq k} \|W_i - x_\varepsilon\| \geq a$ and $|h(x_\varepsilon) - M| \leq \varepsilon$. Let $\varepsilon_n = 1/n$. Then the infinite sequence x_{ε_n}, which belongs to the compact set $[0, 2\pi]^{m-1}$, has a convergent subsequence with some limit x^*. By construction, $x^* \neq W_i$ for any i, and the continuity of h implies that $h(x^*) = M$. This contradiction proves (6.3).

Equation (6.3) implies that, for sufficiently small ε, $S(\varepsilon)$-neighborhoods of points $V_1, \dots, V_k$ have empty intersection. Hence, by definitions (6.1) and (6.2), for sufficiently small ε the following equality is valid:

$$\mathbb{P}\{h(\beta) \geq M - \varepsilon\} = \sum_{i=1}^{k} \mathbb{P}\{h(\beta) \geq M - \varepsilon, \|W_i - \beta\| \leq S(\varepsilon)\}. \tag{6.4}$$

Let us fix some $i \in \{1, \dots, k\}$. Assume that the following event happens for some $\varepsilon > 0$:

$$h(\beta) = h(\beta_1, \dots, \beta_{m-1}) \geq M - \varepsilon, \ \|W_i - \beta\| \leq S(\varepsilon). \tag{6.5}$$

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

By (6.3) there exists $\varepsilon_0 > 0$ such that $S(\varepsilon_0) < \delta/2$, where δ is the number from condition A5. The function $S(\varepsilon)$ is non-decreasing, so for any positive $\varepsilon < \varepsilon_0$ we have

$$S(\varepsilon) < \frac{\delta}{2}. \tag{6.6}$$

Below we deal with $\varepsilon < \varepsilon_0$ only. Since the function h is three-times continuously differentiable in the δ-neighborhood of any maximal point, in this neighborhood we consider the Taylor expansion of the function h at the point W_i with the third-order remainder. For this purpose we introduce the notation

$$\alpha_j = \beta_j - W_i^j \quad \text{and} \quad \alpha = (\alpha_1, \dots, \alpha_{m-1}). \tag{6.7}$$

It is clear that

$$\|\alpha\| = \|\beta - W_i\| < \frac{\delta}{2}.$$

Here α is an element of $\mathbb{R}^{m-1}$ which is considered as a difference of two elements of $\mathbb{R}^{m-1}$ and not as the difference of two sets of angles. By (6.3) and condition A4 it is the same for small ε.

We write the Taylor expansion of the function h at the point W_i. Then we have

$$\begin{aligned} h(\beta) &= h(\beta_1, \dots, \beta_{m-1}) \\ &= h\big(W_i^1 + \alpha_1, W_i^2 + \alpha_2, \dots, W_i^{m-1} + \alpha_{m-1}\big) \\ &= h(W_i) + \sum_{j=1}^{m-1} \frac{\partial h(W_i)}{\partial x_j} \alpha_j + \sum_{1 \le l,s \le m-1} \frac{1}{2} \frac{\partial^2 h(W_i)}{\partial x_l \partial x_s} \alpha_l \alpha_s \\ &\quad + \sum_{1 \le l,s,t \le m-1} \frac{1}{6} \frac{\partial^3 h(W_i + r_{(l,s,t)})}{\partial x_l \partial x_s \partial x_t} \alpha_l \alpha_s \alpha_t, \end{aligned} \tag{6.8}$$

where $r_{(l,s,t)} = c_{(l,s,t)} \cdot (\alpha_1, \dots, \alpha_{m-1})$, and $c_{(l,s,t)} \in (0, 1)$ are constants depending on indices l, s, t and on the function h. According to condition A4, W_i does not lie on the boundary of the definition domain of the continuous function h, so $\partial h(W_i)/\partial x_j = 0$ for all $j \in \{1, \dots, m-1\}$. Hence the linear term in expansion (6.8) is equal to 0.

Consider the matrix

$$A^i = \frac{1}{2} G_i, \tag{6.9}$$

where the G_i are the same as in condition A6. It is clear that the coefficient before $\alpha_l \alpha_s$ in (6.8) is $a_{l,s}^i$ (the element of the matrix A^i). Thus

$$h(\beta) = M + \sum_{1 \le l,s \le m-1} a_{l,s}^i \alpha_l \alpha_s + \frac{1}{6} \sum_{1 \le l,s,t \le m-1} \frac{\partial^3 h(W_i + r_{(l,s,t)})}{\partial x_l \partial x_s \partial x_t} \alpha_l \alpha_s \alpha_t.$$

Therefore condition (6.5) is equivalent to

$$-\sum_{1 \le l,s \le m-1} a_{l,s}^i \alpha_l \alpha_s - \frac{1}{6} \sum_{1 \le l,s,t \le m-1} \frac{\partial^3 h(W_i + r_{(l,s,t)})}{\partial x_l \partial x_s \partial x_t} \alpha_l \alpha_s \alpha_t \le \varepsilon. \tag{6.10}$$

Under conditions (6.5) and (6.6), we estimate the third-order terms in this formula. Since the functions

$$\frac{\partial^3 h(W_i + r)}{\partial x_l \partial x_s \partial x_t}$$

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

are continuous for $|r| \le \delta/2$, there exists $L_1 > 0$ such that

$$\left| \frac{\partial^3 h(W_i + r)}{\partial x_l \partial x_s \partial x_t} \right|$$

does not exceed L_1 for all $|r| \le \delta/2$. Therefore the following inequality holds true:

$$\left| \frac{1}{6} \frac{\partial^3 h(W_i + r_{(l,s,t)})}{\partial x_l \partial x_s \partial x_t} \alpha_l \alpha_s \alpha_t \right| \le L_1 |\alpha_l \alpha_s \alpha_t| \le L_1 \frac{|\alpha_l|^3 + |\alpha_s|^3 + |\alpha_t|^3}{3} \le L_1 \frac{\alpha_l^2 + \alpha_s^2 + \alpha_t^2}{3} S(\varepsilon). \tag{6.11}$$

The last inequality follows from $\|\alpha\| < S(\varepsilon)$. Summing (6.11) over all triplets (l, s, t), we get the inequality

$$\left| \frac{1}{6} \sum_{1 \le l,s,t \le m-1} \frac{\partial^3 h(W_i + r_{(l,s,t)})}{\partial x_l \partial x_s \partial x_t} \alpha_l \alpha_s \alpha_t \right| \le m^2 L_1 S(\varepsilon) \sum_{s=1}^{m-1} \alpha_s^2 = L_2 S(\varepsilon) \sum_{s=1}^{m-1} \alpha_s^2.$$

Therefore the following estimate is valid for all $\|\alpha\| < S(\varepsilon)$:

$$\begin{aligned} L_2 S(\varepsilon) \sum_{s=1}^{m-1} \alpha_s^2 - \sum_{1 \le l,s \le m-1} a_{l,s}^i \, \alpha_l \alpha_s &\ge M - h(W_i + \alpha) \\ &\ge -L_2 S(\varepsilon) \sum_{s=1}^{m-1} \alpha_s^2 - \sum_{1 \le l,s \le m-1} a_{l,s}^i \, \alpha_l \alpha_s. \end{aligned} \tag{6.12}$$

Collecting the results of (6.5), (6.10), and (6.12), we obtain

$$\begin{aligned} &\mathbb{P}\left\{ \left(- \sum_{1 \le l,s \le m-1} a_{l,s}^i \alpha_l \alpha_s - L_2 S(\varepsilon) \sum_{s=1}^{m-1} \alpha_s^2 \right) \le \varepsilon, \ \|\alpha\| < S(\varepsilon) \right\} \\ &\quad \ge \mathbb{P}\{h(\beta) > M - \varepsilon, \ \|W_i - \beta\| \le S(\varepsilon)\} \\ &\quad \ge \mathbb{P}\left\{ \left(- \sum_{1 \le l,s \le m-1} a_{l,s}^i \alpha_l \alpha_s + L_2 S(\varepsilon) \sum_{s=1}^{m-1} \alpha_s^2 \right) \le \varepsilon, \ \|\alpha\| < S(\varepsilon) \right\}. \end{aligned} \tag{6.13}$$

Denote

$$A^i(\varepsilon) = \begin{cases} A^i + L_2 S(\varepsilon) I_{m-1} & \text{for } \varepsilon \ge 0, \\ A^i - L_2 S(-\varepsilon) I_{m-1} & \text{for } \varepsilon \le 0, \end{cases}$$

where A^i is the same as in (6.9), and I_{m-1} is the identity matrix of size $(m-1) \times (m-1)$. Then inequality (6.13) may be rewritten using the scalar product $\langle \cdot, \cdot \rangle$ as

$$\begin{aligned} \mathbb{P}\{-\langle A^i(\varepsilon)\alpha, \alpha\rangle \le \varepsilon, \ \|\alpha\| < S(\varepsilon)\} &\ge \mathbb{P}\{h(\beta) \ge M - \varepsilon, \ \|W_i - \beta\| \le S(\varepsilon)\} \\ &\ge \mathbb{P}\{-\langle A^i(-\varepsilon)\alpha, \alpha\rangle \le \varepsilon, \ \|\alpha\| < S(\varepsilon)\}. \end{aligned} \tag{6.14}$$

Next we need the following lemma.

Lemma 6.1. *There exist constants $\varepsilon_1 > 0$ and $\lambda > 0$ such that for any $|\varepsilon| < \varepsilon_1$ the matrix $A^i(\varepsilon)$ is negative definite, and none of its eigenvalues exceed $-\lambda$.*

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

Proof. It is well known that a symmetric real-valued matrix B is negative definite if and only if all the eigenvalues of the matrix B are negative (see e.g. [5, §4.1, p. 231, Th. 4.1.10]). Therefore, to prove Lemma 6.1, it is enough to show that none of the eigenvalues of the matrix $A^i(\varepsilon)$ exceed some $-\gamma < 0$, for small ε. Let

$$\lambda_1(B) \leq \ldots \leq \lambda_s(B) \tag{6.15}$$

denote the eigenvalues of the Hermitian matrix B of size $s \times s$. The matrix A^i is negative semidefinite (see e.g. [21, §4.5, p. 563]). By condition A6 and (6.9), $\det(A^i) \neq 0$, hence the matrix A^i is negative definite. Using (6.15) it is equivalent to $\lambda_{m-1}(A^i) < 0$.

In what follows, we will need the Weyl theorem formulated below. It may be found in [5, §4.3, p. 239, Th. 4.3.1], for example.

Theorem 6.2. (Weyl.) *Let A, B be Hermitian matrices of size $s \times s$. Then, using (6.15), we have*

$$\lambda_i(A+B) \leq \lambda_{i+j}(A) + \lambda_{s-j}(B)$$

for all $i \in \{1, \ldots, s\}$, $j \in \{0, \ldots, s-i\}$.

Due to the Weyl theorem the eigenvalues of $A^i(\varepsilon)$ do not exceed $\lambda_{m-1}(A^i) + L_2 S(|\varepsilon|)$. Therefore we can put

$$\gamma = \frac{|\lambda_{m-1}(A^i)|}{2}$$

and ε_0 such that

$$S(\varepsilon_0) < \frac{|\lambda_{m-1}(A^i)|}{2L_2}.$$

By (6.3) such γ and ε_0 satisfy the conditions of Lemma 6.1. □

Remark 6.1. Obviously,

$$\lim_{\varepsilon \to 0} \det(A^i(\varepsilon)) = \det(A^i) > 0.$$

Lemma 6.2. *There exist constants $C, D > 0$ such that if, for some $\varepsilon > 0$ and β, the conditions*

(i) $h(\beta) > M - \varepsilon$,

(ii) $\|\alpha\| = \|W_i - \beta\| < S(\varepsilon)$,

(iii) $\varepsilon < C$,

are satisfied, then $\|\alpha\| \leq D\sqrt{\varepsilon}$. The notation α and β are the same as above.

Proof. Let $C = \min(\varepsilon_1, \varepsilon_0)$, where ε_1 is from Lemma 6.1 and ε_0 is from (6.6). According to the Rayleigh theorem (see e.g. [5, §4.2, p. 234, Th. 4.2.2]) the inequality $\langle Bx, x\rangle \leq \lambda_s(B)\|x\|^2$, where $\lambda_s(B)$ is from (6.15), holds for every Hermitian matrix B of size $s \times s$. Then for any $|\varepsilon| < C$ we have $\langle A^i(\varepsilon)x, x\rangle \leq -\gamma\|x\|^2$, where γ is also from Lemma 6.1. Note that by (6.12) the following inequality is valid for positive ε:

$$M + \langle A^i(\varepsilon)\alpha, \alpha\rangle \geq h(W_i + \alpha) > M - \varepsilon.$$

Therefore $\varepsilon \geq -\langle A^i(\varepsilon)\alpha, \alpha\rangle \geq \gamma\|\alpha\|^2$. Hence $\|\alpha\|$ is less than $\sqrt{\varepsilon}/\sqrt{\gamma}$. □

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

The next corollary follows from the proof of Lemma 6.2.

Corollary 6.1. *If* $\varepsilon \geq -\langle A^i(\varepsilon)\alpha, \alpha\rangle$ *and* $0 < \varepsilon < C$, *then* $\|\alpha\| \leq D\sqrt{\varepsilon}$.

Lemma 6.2 completely proves assertion (i) of Theorem 6.1. It remains to prove the second one. For this purpose we need to calculate

$$\mathbb{P}\{\langle -A^i(\pm\varepsilon)\alpha, \alpha\rangle \leq \varepsilon,\ \|\alpha\| < S(\varepsilon)\}$$

for small ε. By Corollary 6.1 condition $\|\alpha\| < S(\varepsilon)$, where $S(\varepsilon)$ is defined by (6.3), follows from the inequality $\langle -A^i(\pm\varepsilon)\alpha, \alpha\rangle \leq \varepsilon$ for sufficiently small ε. Therefore

$$\mathbb{P}\{\langle -A^i(\pm\varepsilon)\alpha, \alpha\rangle \leq \varepsilon,\ \|\alpha\| < S(\varepsilon)\} = \mathbb{P}\{\langle -A^i(\pm\varepsilon)\alpha, \alpha\rangle \leq \varepsilon\}.$$

Below we assume that $0 < \varepsilon < C$, where C is the same as in Lemma 6.2.

Lemma 6.3. *For sufficiently small* ε *there exists* $x^*(\varepsilon) \in [0, 2\pi]^{m-1}$ *such that the following equality holds:*

$$\mathbb{P}\{\langle -A^i(\pm\varepsilon)\alpha, \alpha\rangle \leq \varepsilon\} = \frac{(\pi\varepsilon)^{(m-1)/2}}{\Gamma\left(\frac{m+1}{2}\right)\sqrt{\det(-A^i(\varepsilon))}} \cdot \int_0^{2\pi} \left[p(y) \prod_{l=1}^{m-1} p\big(y + W_i^l + x_l^*(\varepsilon)\big) \right] \mathrm{d}y,$$

where $\|x^*(\varepsilon)\| \leq D\sqrt{\varepsilon}$, *and the constant* D *is introduced in Lemma* 6.2.

Proof. In order to simplify the formulas, we put

$$B = -A^i(\varepsilon) = (b_{i,j})_{i,j=1}^{m-1,m-1}.$$

We also introduce $\beta_0 = \angle xOV_1$ to be the angle between the axis Ox and the vector OV_1, taken counterclockwise. We have

$$\begin{aligned}\mathbb{P}\{\langle -A^i(\pm\varepsilon)\alpha, \alpha\rangle \leq \varepsilon\} &= \mathbb{P}\{\langle B\alpha, \alpha\rangle \leq \varepsilon\} = \mathbb{P}\left\{ \sum_{1\leq s,t\leq m-1} b_{s,t}\alpha_s\alpha_t \leq \varepsilon \right\} \\ &= \int_0^{2\pi} p(y) \int_{\mathbb{R}^{m-1}} \mathbf{1}\left\{ \sum_{1\leq s,t\leq m-1} b_{s,t}x_s x_t \leq \varepsilon \right\} \prod_{l=1}^{m-1} \rho_l(x_l \mid \beta_0 = y)\, \mathrm{d}x_1 \cdots \mathrm{d}x_{m-1}\, \mathrm{d}y, \quad (6.16)\end{aligned}$$

where $\rho_l(x_l \mid \beta_0 = y)$ is the conditional density of $\alpha_l = x_l$ given $\beta_0 = y$. Taking into account that $V_1, \ldots, V_m$ are independent random variables and using (6.7), we obtain $\rho_l(x_l \mid \beta_0 = y) = p\big(y + W_i^l + x_l\big)$.

Also note that in the general case the second integral is taken not over $\mathbb{R}^{m-1}$ but over quotient space $\mathbb{R}^{m-1}/_\sim$, where $x, y \in \mathbb{R}^{m-1}$ and $x \sim y$ if, for any $i \in \{1, \ldots, m-1\}$, we have $x^i - y^i = 2\pi r$ for some $r \in \mathbb{Z}$. The reason is that the values of α which differ by $2\pi r$ correspond to the same angle β. But by Corollary 6.1, inequality $\langle B\alpha, \alpha\rangle \leq \varepsilon$ implies $\|\alpha\| \leq D\sqrt{\varepsilon}$. Therefore, for small ε, all the $\alpha \in \mathbb{R}^{m-1}$ satisfying this inequality correspond to different β.

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

Hence, for sufficiently small ε, we may indeed integrate over $\mathbb{R}^{m-1}$. Using these facts, we continue (6.16):

$$\begin{aligned}
&\mathbb{P}\{\langle -A^i(\pm\varepsilon)\alpha, \alpha\rangle \le \varepsilon\} \\
&\quad = \int_0^{2\pi} p(y) \int_{\mathbb{R}^{m-1}} \mathbf{1}\Big\{ \sum_{1\le s,t\le m-1} b_{s,t} x_s x_t \le \varepsilon \Big\} \prod_{l=1}^{m-1} p\big(y + W_i^l + x_l\big) \mathrm{d}x_1 \cdots \mathrm{d}x_{m-1}\, \mathrm{d}y \\
&\quad = \int_{\mathbb{R}^{m-1}} \mathbf{1}\Big\{ \sum_{1\le s,t\le m-1} b_{s,t} x_s x_t \le \varepsilon \Big\} \int_0^{2\pi} p(y) \prod_{l=1}^{m-1} p\big(y + W_i^l + x_l\big) \mathrm{d}y\, \mathrm{d}x_1 \cdots \mathrm{d}x_{m-1}. \qquad (6.17)
\end{aligned}$$

This formula is obtained just by switching the integration order. In order to calculate this integral we define the set

$$T = \Big\{ x \in \mathbb{R}^{m-1} : \sum_{1\le s,t\le m-1} b_{s,t} x_s x_t \le \varepsilon \Big\} = \{ x \in \mathbb{R}^{m-1} : \langle x, Bx\rangle \le \varepsilon \}. \qquad (6.18)$$

The set T is an ellipsoid with center 0 and configuration matrix $(1/\varepsilon \cdot B)^{-1}$; see [8, p. 97]. It is well known that the ellipsoid with center 0 and symmetric positive semidefinite configuration matrix Q is defined by $\{x \in \mathbb{R}^p : \langle x, Q^{-1}x\rangle \le 1\}$ and its volume is equal to

$$\frac{\pi^{p/2}}{\Gamma(p/2+1)\sqrt{\det Q}}.$$

It is mentioned in [8, p. 103], for example. Therefore the Lebesgue measure of the set T satisfies the equality

$$\mathrm{meas}(T) = \frac{(\varepsilon\pi)^{(m-1)/2}}{\sqrt{\det(B)}\,\Gamma\big(\frac{m+1}{2}\big)}. \qquad (6.19)$$

In what follows, we will need a mean value theorem (see e.g. [12]). Using this theorem we obtain the equality

$$\begin{aligned}
&\int_{\mathbb{R}^{m-1}} \mathbf{1}\Big\{ \sum_{1\le s,t\le m-1} b_{s,t} x_s x_t \le \varepsilon \Big\} \int_0^{2\pi} p(y) \prod_{l=1}^{m-1} p\big(y + W_l^i + x_l\big) \mathrm{d}y\ \mathrm{d}x_1 \cdots \mathrm{d}x_{m-1} \\
&\quad = \int_0^{2\pi} p(y) \prod_{l=1}^{m-1} p\big(y + W_l^i + x_l^*(\varepsilon)\big) \mathrm{d}y \cdot \mathrm{meas}(T), \qquad \text{where } x^*(\varepsilon) \in T. \qquad (6.20)
\end{aligned}$$

Recall that the set T is defined by (6.18). The condition $x^*(\varepsilon) \in T$ together with (6.18) implies

$$\sum_{1\le s,t\le m-1} b_{s,t} x_s^*(\varepsilon) x_t^*(\varepsilon) \le \varepsilon.$$

Therefore Corollary 6.1 implies inequality $\|x^*(\varepsilon)\| \le D\sqrt{\varepsilon}$, where the constant D is introduced in Lemma 6.2. Hence, collecting formulas (6.17), (6.19), (6.20), we finish the proof. □

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

Substituting the result of Lemma 6.3 into inequality (6.14), for small ε we get the relation

$$\begin{aligned}
&\frac{(\pi\varepsilon)^{(m-1)/2}}{\Gamma\left(\frac{m+1}{2}\right)\sqrt{\det(-A^i(\varepsilon))}}\cdot\int_0^{2\pi}\left[p(y)\prod_{l=1}^{m-1}p\big(y+W_i^l+x_l^*(\varepsilon)\big)\right]\mathrm{d}y\\
&\quad\geq \mathbb{P}\{h(\beta)\geq M-\varepsilon,\ \|W_i-\beta\|\leq S(\varepsilon)\}\\
&\quad\geq \frac{(\pi\varepsilon)^{(m-1)/2}}{\Gamma\left(\frac{m+1}{2}\right)\sqrt{\det(-A^i(-\varepsilon))}}\cdot\int_0^{2\pi}\left[p(y)\prod_{l=1}^{m-1}p\big(y+W_i^l+x_l^*(-\varepsilon)\big)\right]\mathrm{d}y.
\end{aligned}$$

Dividing both sides of this inequality by $\varepsilon^{(m-1)/2}$ and tending ε to 0, we get

$$\begin{aligned}
&\lim_{\varepsilon\to 0+}\frac{\pi^{(m-1)/2}}{\Gamma\left(\frac{m+1}{2}\right)\sqrt{\det(-A^i(\varepsilon))}}\cdot\int_0^{2\pi}\left[p(y)\prod_{l=1}^{m-1}p\big(y+W_i^l+x_l^*(\varepsilon)\big)\right]\mathrm{d}y\\
&\quad\geq \lim_{\varepsilon\to 0+}\varepsilon^{-(m-1)/2}\mathbb{P}\{h(\beta)\geq M-\varepsilon,\ \|W_i-\beta\|\leq S(\varepsilon)\}\\
&\quad\geq \lim_{\varepsilon\to 0+}\frac{\pi^{(m-1)/2}}{\Gamma\left(\frac{m+1}{2}\right)\sqrt{\det(-A^i(-\varepsilon))}}\cdot\int_0^{2\pi}\left[p(y)\prod_{l=1}^{m-1}p\big(y+W_i^l+x_l^*(-\varepsilon)\big)\right]\mathrm{d}\beta_0. \qquad (6.21)
\end{aligned}$$

Using

$$\lim_{\varepsilon\to 0}\det(A^i(\varepsilon))=\det(A^i),$$

$$\lim_{\varepsilon\to 0}\int_0^{2\pi}\left[p(y)\prod_{l=1}^{m-1}p\big(y+W_i^l+x_l^*(\varepsilon)\big)\right]\mathrm{d}y=\int_0^{2\pi}\left[p(y)\prod_{l=1}^{m-1}p\big(y+W_i^l\big)\right]\mathrm{d}y,$$

we pass to the limit in (6.21) and obtain

$$\begin{aligned}
&\lim_{\varepsilon\to 0+}\varepsilon^{-(m-1)/2}\mathbb{P}\{h(\beta)\geq M-\varepsilon,\ \|W_i-\beta\|\leq S(\varepsilon)\}\\
&\quad=\frac{\pi^{(m-1)/2}}{\Gamma\left(\frac{m+1}{2}\right)\sqrt{\det(-A^i)}}\cdot\int_0^{2\pi}\left[p(y)\prod_{l=1}^{m-1}p\big(y+W_i^l\big)\right]\mathrm{d}y. \qquad (6.22)
\end{aligned}$$

From (6.9) we may conclude that

$$\det(-A^i)=\frac{\det(-G_i)}{2^{m-1}}. \qquad (6.23)$$

Now, using (6.4), (6.22), and (6.23), we obtain

$$\begin{aligned}
\lim_{\varepsilon\to 0+}\varepsilon^{-(m-1)/2}\mathbb{P}\{h(\beta)\geq M-\varepsilon\}&=\frac{\pi^{(m-1)/2}}{\Gamma\left(\frac{m+1}{2}\right)}\sum_{i=1}^{k}\frac{\int_0^{2\pi}\left[p(y)\prod_{l=1}^{m-1}p\big(y+W_i^l\big)\right]\mathrm{d}y}{\sqrt{\det(-A^i)}}\\
&=\frac{(2\pi)^{(m-1)/2}}{\Gamma\left(\frac{m+1}{2}\right)}\sum_{i=1}^{k}\frac{\int_0^{2\pi}\left[p(y)\prod_{l=1}^{m-1}p\big(y+W_i^l\big)\right]\mathrm{d}y}{\sqrt{\det(-G_i)}},
\end{aligned}$$

and Theorem 6.1 is proved. □

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

7. Proof of Theorem 3.1: part 2

In this section we prove the second part of Theorem 3.1.

Theorem 7.1. *Suppose that a kernel f and points $V_1, \dots, V_n$ satisfy conditions* A *and* B*, which are listed in Section 2. Let M_n be the U-*max *statistic constructed via kernel f, that is, $M_n = \max_{1\le i_1<\dots<i_m\le n} f(V_{i_1}, \dots, V_{i_m})$. Then for every $t > 0$ the following relation holds true:*

$$\lim_{n\to\infty} \mathbb{P}\big\{n^{2m/(m-1)}(M - M_n) \le t)\big\} = 1 - \exp\left(-\frac{t^{(m-1)/2}K_m}{m!}\right),$$

where the constant K_m is introduced in condition (ii) of Theorem 6.1. The rate of convergence is $O\big(n^{-1/(m-1)}\big)$.

Before proving Theorem 7.1, let us mention two theorems that play a key role in studying the limit behavior of U-max statistics. The first theorem is taken from [10], where Lao and Mayer applied the Poisson approximation from the monograph [2]. This is still the main research method in this field.

Theorem 7.2. ([10].) *Let $\xi_1, \xi_2, \dots, \xi_n$ be a sequence of independent identically distributed random elements taking values in a measurable space $(\mathfrak{X}, \mathfrak{A})$, and let the function $f(x_1, \dots, x_m)$ be a real-valued symmetric Borel function, $f\colon \mathfrak{X}^m \to \mathbb{R}$. Let $M_n = \max_J h(\xi_{i_1}, \dots, \xi_{i_m})$ be the U-*max *statistic introduced in (1.2), and define, for any $z \in \mathbb{R}$, the following quantities:*

$$p_{n,z} = \mathbb{P}\{f(\xi_1, \dots, \xi_m) > z\}, \qquad \lambda_{n,z} = \binom{n}{m} p_{n,z},$$
$$\tau_{n,z}(r) = \frac{\mathbb{P}\{f(\xi_1, \dots, \xi_m) > z, f(\xi_{1+m-r}, \xi_{2+m-r}, \dots, \xi_{2m-r}) > z\}}{p_{n,z}}.$$

Then, for all $n \ge m$ and for each $z \in \mathbb{R}$, we have

$$\begin{aligned}&|\mathbb{P}(M_n \le z) - \mathrm{e}^{-\lambda_{n,z}}| \\ &\quad \le \big(1 - \mathrm{e}^{-\lambda_{n,z}}\big) \cdot \left[p_{n,z}\left(\binom{n}{m} - \binom{n-m}{m}\right) + \sum_{r=1}^{m-1} \binom{m}{r}\binom{n-m}{m-r} \tau_{n,z}(r)\right].\end{aligned} \tag{7.1}$$

Remark 7.1. ([10].) If the sample size n tends to infinity, then the right-hand side of (7.1) is of asymptotic order

$$\mathrm{O}\left(p_{n,z}n^{m-1} + \sum_{r=1}^{m-1} \tau_{n,z}(r)n^{m-r}\right),$$

where for $m > 1$ the first term is negligibly small with respect to the sum.

Silverman and Brown [16] have found the conditions for a general theorem used in [2] yielding a non-trivial Weibull law in the limit.

Theorem 7.3. ([16].) *Let the conditions of Theorem 7.2 be satisfied. If, for some sequence of transformations $z_n\colon T \to \mathbb{R}$, $T \subset \mathbb{R}$, the equalities*

$$\lim_{n\to\infty} \lambda_{n,z_n(t)} = \lambda_t > 0, \tag{7.2}$$

$$\lim_{n\to\infty} n^{2m-1} p_{n,z_n(t)} \tau_{n,z_n(t)}(m-1) = 0 \tag{7.3}$$

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

hold, for each $t \in T$, then

$$\lim_{n\to\infty} \mathbb{P}(M_n \le z_n(t)) = \mathrm{e}^{-\lambda_t} \tag{7.4}$$

for all $t \in T$.

Remark 7.2. ([10].) Condition (7.2) implies $p_{n,z} = \mathrm{O}(n^{-m})$. Therefore, according to Remark 7.1, the rate of convergence in (7.4) is

$$\mathrm{O}\left(n^{-1} + \sum_{r=1}^{m-1} n^{2m-r} p_{n,z}\tau_{n,z}(r)\right).$$

Hence, for $m \ge 2$, condition (7.3) can be replaced by

$$\lim_{n\to\infty} n^{2m-r} p_{n,z}\tau_{n,z}(r) = 0 \quad \text{for any } r \in \{1, \dots, m-1\}. \tag{7.5}$$

Now we will use the above assertions to prove Theorem 7.1.

Proof. For any $t > 0$ we define the transformation

$$z_n(t) = M - tn^{-2m/(m-1)}.$$

Let us consider $\lambda_{n,z_n(t)}$ defined in Theorem 7.2. Then

$$\lambda_{n,z_n(t)} = \frac{n!}{m!(n-m)!}\mathbb{P}\{f(V_1, \dots, V_m) > z_n(t)\}.$$

We put $\varepsilon = tn^{-2m/(m-1)}$, then $n^m\varepsilon^{(m-1)/2} = t^{(m-1)/2}$. Let us prove the fulfillment of condition (7.2) of Theorem 7.3. We write

$$\begin{aligned}
\lim_{n\to\infty} \lambda_{n,z_n(t)} &= \lim_{n\to\infty} \frac{n!}{m!(n-m)!}\mathbb{P}\{f(V_1, \dots, V_m) > z_n(t)\} \\
&= \frac{1}{m!} \lim_{n\to\infty} \frac{n!}{n^m(n-m)!} n^m \varepsilon^{(m-1)/2}\varepsilon^{-(m-1)/2}\mathbb{P}\{f(V_1, \dots, V_m) > M - \varepsilon\} \\
&= \frac{1}{m!} t^{(m-1)/2} \lim_{n\to\infty} \left(tn^{-2m/(m-1)}\right)^{-(m-1)/2}\mathbb{P}\left\{f(V_1, \dots, V_m) > M - tn^{-2m/(m-1)}\right\} \\
&= \frac{t^{(m-1)/2}K_m}{m!} \\
&=: \lambda_t > 0.
\end{aligned}$$

In the last line we used statement (ii) of Theorem 6.1. Now we will prove condition (7.5) of Remark 7.2. We formulate this statement as a separate lemma.

Lemma 7.1. *For each $r \in \{1, \dots, m-1\}$ we have the following relation:*

$$\lim_{n\to\infty} n^{2m-r}\mathbb{P}\{f(V_1, \dots, V_m) > z_n(t), f(V_{1+m-r}, \dots, V_{2m-r}) > z_n(t)\} = 0.$$

Proof. Let us introduce the following notation: $\beta_i = \angle V_1 O V_{i+1}$ for $i \in \{1, \dots, 2m-r-1\}$, $\gamma_i = \angle V_{m-r+1} O V_{i+1}$ for $i \in \{m-r, \dots, 2m-r-1\}$. Such notation corresponds to (2.1) and (2.2) for each $i \in \{1, \dots, m-1\}$. It is clear that $\gamma_i = (\beta_i - \beta_{m-r}) \mod 2\pi$ for any $i \ge m$. We introduce the events

$$Q_{i,j} = \left\{\|W_i - (\beta_1, \dots, \beta_{m-1})\| \le D\sqrt{\varepsilon},\ \|W_j - (\gamma_{m-r}, \dots, \gamma_{2m-r-1})\| \le D\sqrt{\varepsilon}\right\},$$

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

where W_i is the same as in condition A4 and the constant D is introduced in Theorem 6.1. It follows from Lemma 6.2 that for small $z_n(t)$ the following equality holds true:

$$\begin{aligned}&\{f(V_1,\dots,V_m)\ge z_n(t)\cap f(V_{1+m-r},\dots,V_{2m-r})\ge z_n(t)\}\\&\quad=\cup_{1\le i,j\le k}\big([f(V_1,\dots,V_m)\ge z_n(t)\cap f(V_{1+m-r},\dots,V_{2m-r})\ge z_n(t)]\cap Q_{i,j}\big).\end{aligned}\tag{7.6}$$

Next we estimate the probability

$$\mathbb{P}\{[f(V_1,\dots,V_m)\ge z_n(t)\cap f(V_{1+m-r},\dots,V_{2m-r})\ge z_n(t)]\cap Q_{i,j}\}.\tag{7.7}$$

By the definition for all elements V_i from $Q_{i,j}$ we have the following bounds for β_i and γ_i: $\|\beta_l-W_i^l\|\le D\sqrt{\varepsilon}$ for each $i<m$, and $\|\gamma_l-W_j^{l-m+r}\|\le D\sqrt{\varepsilon}$.

For $l\ge m$ we obtain

$$\left\|\beta_l-W_i^{m-r}-W_j^{l-m+r}\right\|\le\left\|\beta_l-\beta_{m-r}-W_j^{l-m+r}\right\|+\left\|\beta_{m-r}-W_i^{m-r}\right\|\le 2D\sqrt{\varepsilon}.$$

Let L_p denote the maximal value of density $p(x)$. Using the properties of distribution of β_l, we can estimate the upper bound of probability from (7.7) by

$$\left(2DL_p\sqrt{\varepsilon}\right)^{m-1}\left(4DL_p\sqrt{\varepsilon}\right)^{m-r}.$$

Using formula (7.6) and substituting $\varepsilon=tn^{-2m/(m-1)}$ in the estimate of (7.7), we obtain the inequality

$$\begin{aligned}&n^{2m-r}\mathbb{P}\{f(V_1,\dots,V_m)\ge z_n(t),f(V_{1+m-r},\dots,V_{2m-r})\ge z_n(t)\}\\&\quad\le n^{2m-r}k^2\left(2DL_p\sqrt{tn^{-2m/(m-1)}}\right)^{m-1}\left(4DL_p\sqrt{tn^{-2m/(m-1)}}\right)^{m-r}\\&\quad=\mathrm{O}\big(n^{-(m-r)/(m-1)}\big)\\&\quad=\mathrm{o}(1).\end{aligned}$$

□

Let us return to the proof of Theorem 7.1. Now we may use Theorem 7.3, since all its conditions are verified. Then according to (7.4) we obtain

$$\lim_{n\to\infty}\mathbb{P}(M_n\le z_n(t))=\mathrm{e}^{-\lambda_t}$$

for any $t\in T$. Hence

$$\lim_{n\to\infty}\mathbb{P}\big(M_n\le M-tn^{-2m/(m-1)}\big)=\exp\left(-\frac{t^{(m-1)/2}K_m}{m!}\right).$$

Therefore, for any $t>0$, the following relation is valid:

$$\lim_{n\to\infty}\mathbb{P}\big\{n^{2m/(m-1)}(M-M_n)\le t)\big\}=1-\exp\left(-\frac{t^{(m-1)/2}K_m}{m!}\right).$$

According to Remark 7.2, the convergence rate is

$$\mathrm{O}\left(n^{-1}+\sum_{r=1}^{m-1}p_{n,z_n(t)}\tau_{n,z_n(t)}(r)n^{2m-r}\right).$$

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

From the proof of Lemma 7.1, it follows that

$$\mathrm{O}\big(p_{n,z_n(t)}\tau_{n,z_n(t)}(r)n^{2m-r}\big) = \mathrm{O}\big(n^{-(m-r)/(m-1)}\big),$$

therefore it is $\mathrm{O}\big(n^{-1/(m-1)}\big)$ for $m > 1$. Hence Theorem 7.1 is proved. □

The combination of Theorems 6.1 and 7.1 implies Theorem 3.1.

Acknowledgements

The authors would like to thank Dr Andrei Zaitsev and Dr Dmitry Zaporozhets for their invaluable help concerning this paper. We would like to thank the referees for their careful reading of our manuscript and especially for the helpful comments and suggestions.

Funding information

The work of Ekaterina Simarova was supported by the Ministry of Science and Higher Education of the Russian Federation, agreement no. 075-15-2019-1619. The work of Yakov Nikitin was supported by joint grant RFBR-DFG no. 20-51-12004.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] Alexander, R. and Stolarsky, K. B. (1974). Extremal problems of distance geometry related to energy integrals. *Trans. Amer. Math. Soc.* **193**, 1–31.
[2] Barbour, A. D., Holst, L. and Janson, S. (1992). *Poisson Approximation*. Oxford University Press, London.
[3] Halmos, P. (1946). The theory of unbiased estimation. *Ann. Math. Statist.* **17**, 34–43.
[4] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19**, 293–325.
[5] Horn, R. A. and Johnson, C. R. (2012). *Matrix Analysis*. Cambridge University Press.
[6] Koroleva, E. V. and Nikitin, Ya. Yu. (2014). *U*-max-statistics and limit theorems for perimeters and areas of random polygons. *J. Multivariate Anal.* **127**, 98–111.
[7] Korolyuk, V. S. and V. Borovskikh, Y. (1994). *Theory of U-Statistics*. Springer, Dordrecht.
[8] Kurzhanski, A. and Valyi, I. (1997). *Ellipsoidal Calculus for Estimation and Control*. Birkhäuser, Basel; Springer Science and Business Media, New York.
[9] Lao, W. (2010). Some weak limit laws for the diameter of random point sets in bounded regions. Doctoral thesis, Karlsruhe Institute of Technology.
[10] Lao, W. and Mayer, M. (2008). *U*-max-statistics. *J. Multivariate Anal.* **99**, 2039–2052.
[11] Lee, A. J. (2019). *U-Statistics: Theory and Practice*. Routledge.
[12] Makarov, B. and Podkorytov, A. (2013). *Real Analysis: Measures, Integrals and Applications*. Springer, London.
[13] Mardia, K. V. and Jupp, P. E. (2000). *Directional Statistics*. Wiley.
[14] Mayer, M. (2008). Random diameters and other *U*-max-statistics. Doctoral thesis, University of Bern.
[15] Polevaya, T. A. and Nikitin, Ya. Yu (2019). Limit theorems for areas and perimeters of random inscribed and circumscribed polygons. *Zap. Nauchn. Sem. POMI* **486**, 200–213 (in Russian).
[16] Silverman, F. and Brown, T. (1978). Short distances, flat triangles, and Poisson limits. *J. Appl. Prob.* **15**, 815–825.
[17] Simarova, E. N. (2020). Limit theorems for generalized perimeters of random inscribed polygons I. *Vestnik St. Petersburg Univ. Math. Mech. Astronom.* **7**, 678–687 (in Russian). English translation: *Vestnik St. Petersburg Univ. Math.* **53**, 434–442.

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press

[18] SIMAROVA, E. N. (2021). Limit theorems for generalized perimeters of random inscribed polygons II. *Vestnik St. Petersburg Univ. Math. Mech. Astronom.* **8**, 101–110 (in Russian). English translation: *Vestnik St. Petersburg Univ. Math.* **54**, 78–85.
[19] TOTH, L. F. (1956). On the sum of distances determined by a pointset. *Acta Math. Hungar.* **7**, 397–401.
[20] YAGLOM, I. M. AND BOLTYANSKII, V. (1961). *Convex Figures*. Holt, Rinehart and Winston, New York.
[21] ZORICH, V. A. (2015). *Mathematical Analysis I*. Springer, Berlin and Heidelberg.

https://doi.org/10.1017/jpr.2021.94 Published online by Cambridge University Press