EXTREMAL PROCESSES[1]

By Meyer Dwass

Northwestern University

1. Introduction. Suppose $X_1, X_2, \cdots$ is a sequence of independent and identically distributed random variables, and $M_n = \max(X_1, \cdots, X_n)$, $n = 1, 2, \cdots$. Necessary and sufficient conditions for the existence of sequences of constants $\{a_n\}$ and $\{b_n\}$ such that $\lim_{n\to\infty} P\{(M_n - a_n)/b_n < t\} = F(t)$ should exist for every real t, with F being a non-degenerate c.d.f., are well known. It is also known that up to a location or scale parameter, F must have one of the following three forms:

$$
\begin{aligned}
F_1(t, \lambda) &= 0, && t \leqq 0,\\
&= \exp -\lambda t^{-\alpha}, && t > 0;\\
F_2(t, \lambda) &= \exp -\lambda(-t)^{\alpha}, && t \leqq 0,\\
&= 1, && t > 0;\\
F_3(t, \lambda) &= \exp[-\lambda(\exp -t)], && -\infty < t < \infty;
\end{aligned}
\tag{1.1}
$$

λ being a positive constant, [3]. (In deference to a long-established usage, we wish to point out that the usual notations for F_1, F_2, F_3 are Φ_α, Ψ_α, and Λ respectively.)

The purpose of this paper is to study the stochastic processes, $Y(t)$, which are in a natural sense the "limits," as $n \to \infty$, of the processes

$$Y_n(t) = (M_{[tn]+1} - a_n)/b_n, \qquad 0 \leqq t < \infty.$$

(A study of the passage to the limit is given in the following paper by John Lamperti, [4].) The "limiting" process, $Y(t)$, will be defined rigorously in Section 3. There are, of course, three possible processes, according to whether the X_i's belong to the "laws of attraction" F_1, F_2, or F_3 respectively. We will refer to these processes as *extremal processes* of types 1, 2, 3.

Extremal processes bear a natural analogy with stable processes. Thus, suppose that $U_1, U_2, \cdots$ is a sequence of independent and identically distributed random variables, $S_n = U_1 + \cdots + U_n$, $n = 1, 2, \cdots$, and $\{c_n\}$ and $\{d_n\}$ are sequences of constants such that $\{(S_n - c_n)/d_n\}$ converges in law to a non-degenerate distribution, which is necessarily a distribution of stable type. Then if we define a process

$$Z_n(t) = (S_{[tn]} - c_n)/d_n, \qquad 0 \leqq t < \infty,$$

Received September 17, 1963; revised 25 May 1964.

[1] Research sponsored by Air Force Office of Scientific Research Grant No. AF-AFOSR-67-63.

there is a "limiting" process $Z(t)$ which is in a natural way obtained from $Z_n(t)$ as $n \to \infty$. This $Z(t)$ is a stable process of appropriate type, and is explicitly a process such that the joint distribution of $Z(t_1), \cdots, Z(t_m)$ is exactly the limiting joint distribution of $Z_n(t_1), \cdots, Z_n(t_m)$, as $n \to \infty$, for every choice of positive constants $t_1, \cdots, t_m$, and any positive integer m. This is exactly how extremal processes are being defined, except the role of the sum is now being played by the maximum.

Finally we remark that maxima and sums are only two special cases of functionals which give rise to processes which have interesting structure; it would seem worthwhile to study additional cases of this sort.

2. Preliminary results. The purpose of this section is to give several preliminary results which are of independent interest and will motivate some of the results about $Y(t)$ which will be given in Section 4. As before, we assume that $\{X_n\}$ is a sequence of independent and identically distributed random variables, with *continuous* c.d.f.'s. *However, in this section it is not necessary to assume that* $(M_n - a_n)/b_n$ *converges in law for any choice of constants* $\{a_n\}, \{b_n\}$. We suppose that λ, μ are two constants such that $0 < \lambda < \mu$. Let $N_n(\lambda, \mu)$ denote the number of indices i, among $[n\lambda], [n\lambda] + 1, \cdots, [n\mu]$ for which $X_i = M_i$. In other words, $N_n(\lambda, \mu)$ is the number of new maxima occurring among $X_{[n\lambda]}, \cdots, X_{[n\mu]}$. The assumption of continuous c.d.f.'s guarantees that ties among the X_i's occur with probability zero, and hence with probability one the successive maxima are *strict* maxima. We will also want to refer to the indices at which these maxima occur. We find it convenient to index these by counting from right to left. Thus let $W_{jn\mu}$ be the index of the jth maximum among indices $1, \cdots, [n\mu]$, counting from the right, $(j = 1, \cdots, [n\mu])$.

THEOREM 2.1. *As* $n \to \infty$, (a) $N_n(\lambda, \mu)$ *is asymptotically Poisson distributed with parameter* $\log \frac{\mu}{\lambda}$, (b) $W_{1n\mu}/n\mu, W_{2n\mu}/W_{1n\mu}, \cdots, W_{kn\mu}/W_{(k-1)n\mu}$ *are asymptotically independent and uniformly distributed in* $(0, 1)$.

PROOF. Let $C(i)$ equal 1 if X_i equals M_i, and 0 otherwise. It is an elementary but very useful fact that $C(1), C(2), \cdots$, form a sequence of independent random variables with $EC(i) = 1/i$. (More generally, if $R(i)$ denotes the rank of X_i in the set $X_1, \cdots, X_i$, then $R(1), R(2), \cdots$ form a sequence of independent random variables, $R(i)$ being uniformly distributed on the integers $1, \cdots, i$. This was pointed out by Dwass, Theorem 1 [1], and was discovered again by Rényi [5].) The proof of (a) is now routine using the fact that $N_n(\lambda, \mu) = \sum_{[\lambda n]}^{[\mu n]} C(i)$. Part (b) follows from (a) by a routine argument which we leave to the reader.

REMARK. We should point out that (a) follows easily from (b) by the following elementary fact. Suppose $W_1, W_2, \cdots$ is a sequence of independent random variables each uniformly distributed in $(0, 1)$. Let N be the number of indices i for which $\prod_{j=1}^{i} W_j$ lies in $(t, 1)$, where t is in $(0, 1)$. Then N is Poisson distributed with parameter $-\log t$ by the following argument. Since N equals the number of indices i for which $\sum_{j=1}^{i} (-\log W_j)$ lie in $(0, -\log t)$, and since $\{-\log W_j\}$

forms a sequence of independent, exponential (parameter 1) random variables, then N is Poisson distributed with parameter $-\log t$. Using this fact and (b) it is now easy to conclude that (a) follows.

For the next theorem we define for $u < v$, $J(u, v)$ = number of indices i for which $X_i = M_i$ and $u < X_i < v$. In other words, $J(u, v) + 1$ is the number of new maxima that must be achieved in order to bring M_i from a height below u to a height above v. Denote $P(X_n < t) = G(t)$.

THEOREM 2.2. *Same assumptions as for Theorem* 2.1. *Let u, v be such that G is strictly increasing at u and v and $0 < G(u) < G(v) < 1$. Then $J(u, v)$ is Poisson distributed with parameter $-\log [1 - G(v)]/[1 - G(u)]$.*

PROOF. Consider the random variables $G(X_1)$, $G(X_2)$, $\cdots$. These are independent and uniformly distributed in $(0, 1)$. Maxima occur for the $\{G(X_n)\}$ sequence at exactly the same indices at which they occur for the original $\{X_n\}$ sequence. Moreover, $u < X_n < v$ if and only if $G(u) < G(X_n) < G(v)$. Let N_1, N_2, $\cdots$ be the succession of indices at which the maxima occur, exceeding u, that is $X_i = M_i > u$. It is easy to verify that $G(X_{N_1})$ is uniformly distributed in $(G(u), 1)$, $G(X_{N_2})$ is uniformly distributed in $(G(X_{N_1}), 1)$, etc. Hence, defining $K(x) = -\ln (1 - G(x))$, $K(X_{N_1}) - K(u)$, $\cdots$, $K(X_{N_k}) - K(X_{N_{k-1}})$ form a sequence of independent, exponentially distributed random variables, and the required result is a standard one relating exponential random variables with the Poisson distribution.

REMARKS. It is interesting to compare this theorem with the previous one. Here the result is not an asymptotic one, whereas in Theorem 2.1 the result is asymptotic. On the other hand, the result of Theorem 2.1 does not depend on the form of G, whereas the result of this theorem does.

3. Definition of the extremal processes. In order to simplify somewhat the discussion that follows, we will suppose that X_n is a sequence of independent and identically distributed random variables, whose common c.d.f. is one of the "extreme-value" distributions F_1, F_2, or F_3. Much of what follows would remain the same if we suppose only that the X_n's "belong to the domain of attraction" of these distributions, that is that $(M_n - a_n)/b_n$ converges in law to one of these three distributions for appropriate $\{a_n\}$, $\{b_n\}$.

Corresponding to each of the three possible distributions for the X_n's F_1, F_2, or F_3 as defined in (1.1), we define random processes $Y_{n,1}(t)$, $Y_{n,2}(t)$, or $Y_{n,3}(t)$, as follows: for $0 \leqq t < \infty$,

$$Y_{n,1}(t) = M_{[tn]+1}/n^{1/\alpha}, \qquad Y_{n,2}(t) = M_{[tn]+1}n^{1/\alpha}, \qquad Y_{n,3}(t) = M_{[tn]+1} - \log n. \tag{3.1}$$

The following is easy to verify and we omit the proof.

LEMMA 3.1. *For any constants t_1, $\cdots$, t_k satisfying $0 \leqq t_1 < t_2 < \cdots < t_k$, the asymptotic joint distribution as $n \to \infty$ of the vector $[Y_{n,i}(t_1), \cdots, Y_{n,i}(t_k)]$ is the same as the distribution of a vector $[U_1, \max (U_1, U_2), \cdots, \max (U_1, U_2, \cdots, U_k)]$ where U_1, U_2, $\cdots$, U_k are independent random variables whose* c.d.f.'s *are* $F_i(t, \lambda t_1), F_i(t, \lambda(t_2 - t_1)), \cdots, F_i(t, \lambda(t_k - t_{k-1})), i = 1, 2, 3$. This lemma suggests defining stochastic processes $Y_i(t)$ whose finite dimensional distributions are exactly the limiting finite dimensional distributions of $Y_{n,i}(t)$ described above. We formalize this in the following definition.

DEFINITION. *By an extremal process of type* i ($i = 1, 2, 3$) *we mean a stochastic process* $Y_i(t)$, $0 \leqq t < \infty$, *defined so that for* $0 \leqq t_1 < t_2 < \cdots < t_k$ *the joint distribution of* $Y_i(t_1), Y_i(t_2), \cdots, Y_i(t_k)$ *is the same as the distribution of*

$$U_1, \max(U_1, U_2), \cdots, \max(U_1, U_2, \cdots, U_k),$$

where $U_1, \cdots, U_k$ *are independent random variables whose* c.d.f.'s *are*

$$F_i(t, \lambda t_1), F_i(t, \lambda(t_2 - t_1)), \cdots, F_i(t, \lambda(t_k - t_{k-1})), \qquad i = 1, 2, 3.$$

It is evident from Lemma 3.1 that the above defined processes satisfy the consistency requirements of Kolmogorov, and hence a measure space exists on which such a process can be consistently defined. Moreover, we can suppose, without loss of generality, that we have separable versions of these processes whose sample functions are non-decreasing with probability one. In fact, Theorems 4.2, 4.3, 4.4, below, will give such representations of these processes explicitly.

4. Properties of extremal processes. The first properties to be described are completely expected as they are the limiting analogues of those given in Theorem 2.1. In fact a direct proof based on that theorem and an appropriate "invariance principle" argument should not be too difficult. However, we approach the matter directly. Let $Y(t)$ denote an extremal process of any one of the three types.

LEMMA 4.1. *Suppose* $0 < t_1 < t_2 < \cdots < t_k$. *Define for* $0 < a < b$, $C(a, b) = 1$ *if* $Y(b) > Y(a)$, 0 *otherwise. Then* $C(t_1, t_2), C(t_2, t_3), \cdots, C(t_{k-1}, t_k)$ *are independent random variables. Also,* $EC(a, b) = a/b$.

PROOF. Define $C_n(a, b) = 1$ if $Y_n(b) > Y_n(a)$, 0 otherwise, where $Y(t)$ is one of the processes defined in (3.1) of the same type as $Y(t)$. If $e_1, e_2, \cdots, e_{k-1}$ are each 0 or 1, then

$$(4.1) \qquad \begin{aligned} &P\{(C_n(t_1, t_2) = e_1, \cdots, C_n(t_{k-1}, t_k) = e_k\} \\ &\qquad = P\{C_n(t_1, t_2) = e_1\} \cdot \cdots \cdot P\{C_n(t_{k-1}, t_k) = e_k\} \end{aligned}$$

simply because of the independence of the random variables $\{C(i)\}$ described in the proof of Theorem 2.1. The independence asserted in the theorem will be established if it is permissible to let n go to ∞ in (4.1) and then conclude that the limits of the probabilities equal the same expressions with the n deleted. This is permissible since any of the events in (4.1) are determined by $Y_n(t_1), \cdots, Y_n(t_k)$ whose distribution law converges to that of $Y(t_1), \cdots, Y(t_k)$. Finally, using Lemma 3.1, we conclude that

$$EC(a, b) = P\{Y(b) > Y(a)\} = \int_{-\infty}^{\infty} [1 - F(t, \lambda(b - a))]F(dt, \lambda a)$$

where F is either F_1, F_2 or F_3, according to the corresponding type of $Y(t)$. The integral is easily evaluated to be $1 - a/b$ in each case.

THEOREM 4.1. *If* $0 < a < b$, *then with probability* 1, *any extremal process* $Y(t)$ *is a step function with only a finite number of jumps in* $[a, b]$. *The number of jumps is Poisson distributed with parameter* $\log b/a$.

PROOF. Since $Y(t)$ is separable, and its sample functions are non-decreasing with probability 1, then

$$C(a, a + [b - a]/m) + \cdots + C(b - [b - a]/m, b) \equiv D_m$$

converges with probability one as $m \to \infty$ to the number of discontinuities of $Y(t)$ in $[a, b]$, which number may possibly be ∞. By a routine computation,

$$\lim_{m\to\infty} P(D_m = k) = (\exp - \ln b/a)(\ln b/a)^k/k!.$$

Since the sum of these probabilities for $k = 0, 1, 2, \cdots$ equals 1, there are only a finite number of discontinuities with probability 1 and this number is Poisson distributed with parameter $\log b/a$.

REMARKS. Denoting the ith discontinuity from the right of $Y(t)$ in $(0, a)$ by W_i, $(a > 0)$, then W_1/a, W_2/W_1, $\cdots$, is a sequence of independent random variables, each uniform in $(0, 1)$. Also, the jump points of the process $V(t) = Y(e^t)$ form a regular Poisson process.

Thus far a description has been given only of the discontinuity places in extremal processes. In what follows we want to give a simple representation for extremal processes in terms of certain sequences of independent random variables. This will be analogous to (but more complicated than) the well-known representation of a Poisson process in terms of the successive independent, exponential random variables which represent the waiting times for new jumps.

The following theorems give explicit representations for extremal processes $Y_i(t)$, for t in (s, ∞) where $s > 0$. By a representation we mean a stochastic process whose finite dimensional distributions in (s, ∞) agree with those of the processes $Y_i(t)$ as defined in Section 3. For simplicity, and without loss of generality we suppose hereafter that the parameter λ is equal to 1 throughout.

THEOREM 4.2 (representation of extremal processes of type I). *Let* Y_s, Z_1, Z_2, $\cdots$, V_1, V_2, $\cdots$ *be independent random variables,* Y_s *having the* c.d.f. $F_1(\cdot, s)$, *the* Z_i's *being identically distributed exponential (parameter* $= 1$) *random variables, the* V_i's *being identically distributed and uniform on* $(0, 1)$. *Define the process* $\tilde{Y}_1(t)$, t *in* (s, ∞), $s > 0$, *as follows)*

$$\begin{aligned}
\tilde{Y}_1(t) &= Y_s, && t \text{ in } [s, s + Y_s^\alpha Z_1) \equiv [s, s_1),\\
&= Y_s/V_1^{1/\alpha}, && t \text{ in } [s_1, s_1 + Y_s^\alpha Z_2/V_1) \equiv [s_1, s_2)\\
&= Y_s/(V_1V_2)^{1/\alpha}, && t \text{ in } [s_2, s_2 + Y_s^\alpha Z_3/V_1V_2) \equiv [s_2, s_3),
\end{aligned}$$

etc. Then $\tilde{Y}_1(t)$ *is a representation of* $Y_1(t)$ *for* t *in* (s, ∞).

REMARK. In other words, the process $\tilde{Y}(t)$ evolves as follows. At time s the height is Y_s. The process remains at that height an exponentially distributed

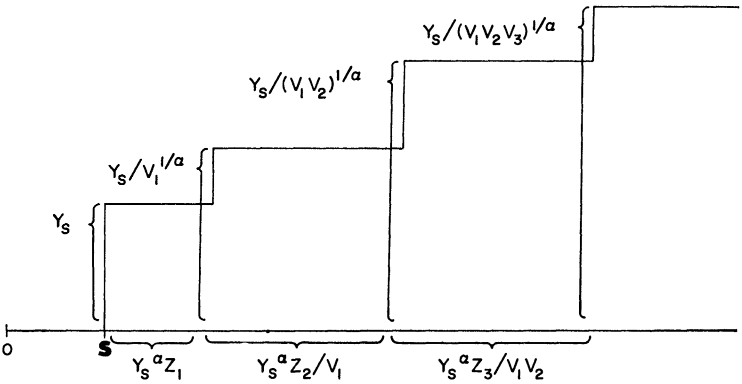

FIG. 1

amount of time, $Y_s^\alpha Z_1$ and then jumps to a new height, $Y_s/V_1^{1/\alpha}$. It remains at that height an exponentially distributed amount of time which is that height raised to the αth power times Z_2 and then jumps to a new height which is the last height divided by $V_2^{1/\alpha}$, etc. Every time the process reaches a new height it remains in that height an exponential amount of time which equals the αth power of that height times the next Z. It then jumps to a new height which is the last height divided by the $(1/\alpha)$th power of the next V. (See Figure 1.)

THEOREM 4.3 (representation of extremal process of type II). *Let Y_s, Z_1, $Z_2, \cdots, V_1, V_2, \cdots$ be independent random variables, Y_s having the* c.d.f. *$F_2(\cdot, s)$, the Z_i's being identically distributed exponential (parameter $= 1$) random variables, the V_i's being identically distributed and uniform on $(0, 1)$. Define the process $\tilde{Y}_2(t)$, t in (s, ∞), $s > 0$, as follows*

$$\begin{aligned} \tilde{Y}_2(t) &= Y_s, && t \text{ in } [s, s + Z_1/(-Y_s)^\alpha) \equiv [s, s_1), \\ &= Y_s V_1^{1/\alpha}, && t \text{ in } [s_1, s_1 + Z_2/(-Y_s)^\alpha V_1) \equiv [s_1, s_2), \\ &= Y_s (V_1 V_2)^{1/\alpha}, && t \text{ in } [s_2, s_2 + Z_3/(-Y_s)^\alpha V_1 V_2) \equiv [s_2, s_3), \end{aligned}$$

etc. Then $\tilde{Y}_2(t)$ is a representation of $Y_2(t)$ for t in (s, ∞).

THEOREM 4.4 (representation of extremal process of type III). *Let Y_s, Z_1, $Z_2, \cdots, V_1, V_2, \cdots$ be independent random variables, Y_s having the* c.d.f. *$F_3(\cdot, s)$, the Z_i's and V_i's being identically distributed exponential (parameter $=$ 1) random variables. Define the process $\tilde{Y}_3(t)$ as follows:*

$$\begin{aligned} \tilde{Y}_3(t) &= Y_s, && t \text{ in } [s, s + e^{Y_s} Z_1) \equiv [s, s_1), \\ &= Y_s + V_1, && t \text{ in } [s_1, s_1 + e^{Y_s + V_1} Z_2) \equiv [s_1, s_2), \\ &= Y_s + V_1 + V_2, && t \text{ in } [s_2, s_2 + e^{Y_s + V_1 + V_2} Z_3) \equiv [s_2, s_3), \end{aligned}$$

etc. Then $\tilde{Y}_3(t)$ is a representation of $Y_3(t)$ for t in (s, ∞).

PROOF OF THEOREMS 4.2, 4.3 and 4.4. It is clear from the definitions of $Y_i(t)$ and $\tilde{Y}_i(t)$ that these are Markov processes with stationary transition functions.

For $\tilde{Y}_i(t)$ this follows from the role of the exponential variables $\{Z_i\}$ in every case. The proof of the theorems will be accomplished by showing that an appropriate version of the infinitesimal generators of the processes $Y_i(t)$ and $\tilde{Y}_i(t)$ coincide. To be specific, it is not hard to verify that these processes satisfy the requirements of Theorem 2 of [2]. Hence it will suffice to show that for any bounded, real-valued, measurable function f,

$$\begin{aligned} &\lim_{t\to 0} E\left\{\frac{f[Y_i(t+s)] - f[Y_i(s)]}{t} \,\middle|\, Y_i(s) = x\right\} \\ &\qquad = \lim_{t\to 0} E\left\{\frac{f[\tilde{Y}_i(t+s)] - f[\tilde{Y}_i(s)]}{t} \,\middle|\, \tilde{Y}_i(s) = x\right\}. \end{aligned} \tag{4.2}$$

(It would actually be sufficient to show this for a smaller class of functions, the class B_0 in Dynkin's notation [2].)

The relevant calculations are given below:

	$E\{f[Y(s+t)] \mid Y(s) = x\}$	$E\{f[\tilde{Y}(s+t)] \mid \tilde{Y}(s) = x\}$
Type I	$f(x) \exp -x^{-\alpha}t +$ $\int_x^\infty \alpha t v^{-(\alpha+1)} \exp -tv^{-\alpha} f(v)\, dv$	$f(x) \exp -x^{-\alpha}t + o(t) +$ $\int_0^1 \int_0^t f(x/v^{1/\alpha}) x^{-\alpha} \exp -(x^{-\alpha})$ $\cdot \exp -[x^{-\alpha}v(t-u)]\, du\, dv$
Type II	$f(x) \exp -(-x)^{\alpha}t +$ $\int_x^0 \alpha t(-v)^{\alpha-1} \exp -t(-v)^{\alpha} f(v)\, dv$	$f(x) \exp -(-x)^{\alpha}t + o(t) +$ $\int_0^1 \int_0^t f(xv^{1/\alpha})(-x)^{\alpha} \exp -(-x)^{\alpha}$ $\cdot \exp -[(-x)^{\alpha}v(t-u)]\, du\, dv$
Type III	$f(x)(\exp -e^{-x}t) +$ $\int_x^\alpha t(\exp -e^{-v}t)(\exp -v) f(v)\, dv$	$f(x) \exp -(e^{-x}t) + o(t) +$ $\int_0^1 \int_0^t f(x+v)(\exp -x)(\exp -e^{-x}u)$ $\cdot \exp -[e^{-x-v}(t-u)]\, du\, dv$

It is now easy to verify that the limits asserted in (4.2) exist and are equal to

$$-f(x)x^{-\alpha} + \alpha \int_x^\infty f(v)v^{-(\alpha+1)}\, dv, \qquad -f(x)(-x)^{\alpha} + \alpha \int_x^0 f(v)(-v)^{\alpha-1}\, dv,$$

and

$$-f(x) \exp - x + \int_0^\infty f(x+v)(\exp - x - v)\, dv$$

for types I, II, III respectively. This completes the proof.

Finally, by means of the representations $\tilde{Y}_i(t)$ we can prove counterparts of Theorem 2.2.

THEOREM 4.5. *Suppose* $s > 0$ *and* $Y_i(s) = u$. *Let* $J(u, v)$, $(u < v)$, *be the number of discontinuities in* $Y_i(t)$, $t > s$, *up to the time for which* $Y_i(t)$ *first exceeds* v. *Then* $J(u, v)$ *is Poisson distributed with parameters indicated by the following table.*

Process type	Restrictions on u, v	Poisson parameter
I	$0 < u < v < \infty$	$\alpha \log (v/u)$
II	$-\infty < u < v < 0$	$\alpha \log (u/v)$
III	$-\infty < u < v < \infty$	$v - u$

PROOF. We give the proof only for case I. The others are similar. Using Theorem 4.2,

$$J(u, v) = \text{last } k \text{ for which } u/(V_1 \cdots V_k)^{1/\alpha} < v$$
$$= \text{last } k \text{ for which } \sum_1^k - \log V_i < \alpha \log (v/u).$$

Since the $-\log V_i$'s are independent and exponential (parameter 1) random variables, then $J(u, v)$ is Poisson distributed with parameter $\alpha \log (v/u)$.

REFERENCES

[1] DWASS, M. (1960). Some k-sample rank order tests. *Contributions to Probability and Statistics* 198–202. Stanford Univ. Press.

[2] DYNKIN, E. B. (1956). Markov processes and semigroups of operators. *Teor. Veroyatnost. i Primenen* **1** 22–33. (English translation by SIAM).

[3] GNEDENKO, B. V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. of Math.* (2) **44** 423–453.

[4] LAMPERTI, JOHN (1964). On extreme order statistics. *Ann. Math. Statist.* **35** 1726–1737.

[5] RÉNYI, A. (1962). Théorie des éléments saillants d'une suite d'observations. *Colloquium on Combinatorial Methods in Probability Theory* 104–117. Mat. Inst., Aarhus Univ