Ann. Inst. Statist. Math.
Vol. 46, No. 2, 343–349 (1994)

ON THE LIMIT BEHAVIOUR OF THE JOINT DISTRIBUTION FUNCTION OF ORDER STATISTICS

H. FINNER AND M. ROTERS

FB IV-Mathematik/Statistik, Universität Trier, D-54286 Trier, Germany

(Received December 11, 1992; revised June 11, 1993)

Abstract. For $k \in \mathbb{N}_0$ fixed we consider the joint distribution function F_n^k of the $n-k$ smallest order statistics of n real-valued independent, identically distributed random variables with arbitrary cumulative distribution function F. The main result of the paper is a complete characterization of the limit behaviour of $F_n^k(x_1, \dots, x_{n-k})$ in terms of the limit behaviour of $n(1-F(x_n))$ if n tends to infinity, i.e., in terms of the limit superior, the limit inferior, and the limit if the latter exists. This characterization can be reformulated equivalently in terms of the limit behaviour of the cumulative distribution function of the $(k+1)$-th largest order statistic. All these results do not require any further knowledge about the underlying distribution function F.

Key words and phrases: Extreme order statistic, extreme value distribution, joint distribution function of order statistics, multiple comparisons, Poisson distribution, uniform distribution.

1. The main result

Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real-valued independent, identically distributed random variables with arbitrary cumulative distribution function (cdf) F and let $X_{1:n} \le \cdots \le X_{n:n}$ denote the order statistics of $X_1, \dots, X_n$. Moreover, let $F_n^k(x_1, \dots, x_{n-k}) = P(X_{1:n} \le x_1, \dots, X_{n-k:n} \le x_{n-k})$ for all $x = (x_1, \dots, x_{n-k}) \in \mathbb{R}^{n-k}$ and $n > k \in \mathbb{N}_0$. For $k \in \mathbb{N}_0$ we define $g_k(c) = \exp(-c)\sum_{j=0}^k c^j/j!$, $c \in [0, \infty)$, and $g_k(\infty) = \lim_{c\to\infty} g_k(c) = 0$, so that g_k can be regarded as a decreasing homeomorphism from $[0, \infty]$ onto $[0, 1]$. Notice that $g_k(c)$, $c \in (0, \infty)$, equals $P_c(Z \le k)$, where Z has a Poisson distribution with parameter c under P_c. The main purpose of the present paper is to prove

THEOREM 1.1. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real-valued independent, identically distributed random variables with cdf F, let $k \in \mathbb{N}_0$, $c \in [0, \infty]$, and $x_n \in (-\infty, \infty]$ such that $\alpha_n = 1 - F(x_n) < 1$ for all $n \in \mathbb{N}$.*

(a)

$$\lim_{n\to\infty} n\alpha_n = c \tag{1.1}$$

implies

$$\lim_{n\to\infty} F_n^k(x_1,\dots,x_{n-k}) = g_k(c). \tag{1.2}$$

(b) *If* $c<\infty$ *or* $(x_n)_{n\in\mathbb{N}}$ *is non-decreasing, then* (1.2) *implies* (1.1).

(c) $\limsup_{n\to\infty} n\alpha_n = c$ *iff* $\liminf_{n\to\infty} F_n^k(x_1,\dots,x_{n-k}) = g_k(c)$.

(d) (i) $\liminf_{n\to\infty} n\alpha_n = c$ *implies* $\limsup_{n\to\infty} F_n^k(x_1,\dots,x_{n-k}) \le g_k(c)$.

(ii) *If, in addition,* $(x_n)_{n\in\mathbb{N}}$ *is a non-decreasing sequence, then* $\liminf_{n\to\infty} n\alpha_n = c$ *iff* $\limsup_{n\to\infty} F_n^k(x_1,\dots,x_{n-k}) = g_k(c)$.

A well-known result related to Theorem 1.1 and concerning the cdf of the $(k+1)$-th largest order statistic is available e.g. in Leadbetter *et al.* ((1983), p. 32), i.e.,

$$\lim_{n\to\infty} n\alpha_n = c \quad \text{iff} \quad \lim_{n\to\infty} P(X_{n-k:n} \le x_{n-k}) = g_k(c) \tag{1.3}$$

for all $c\in[0,\infty]$. As a consequence of (1.3) we also have

$$\liminf_{n\to\infty} n\alpha_n = c \quad \text{iff} \quad \limsup_{n\to\infty} P(X_{n-k:n} \le x_{n-k}) = g_k(c), \tag{1.4}$$

$$\limsup_{n\to\infty} n\alpha_n = c \quad \text{iff} \quad \liminf_{n\to\infty} P(X_{n-k:n} \le x_{n-k}) = g_k(c). \tag{1.5}$$

It is striking that Theorem 1.1 yields the full analogue to (1.3), (1.4) and (1.5) whenever the sequence $(x_n)_{n\in\mathbb{N}}$ is non-decreasing. If this assumption is violated, the equivalence of (1.1) and (1.2) is only valid for $c<\infty$. In Section 3 we present an example showing that for $c=\infty$ (1.2) does not imply (1.1). Moreover, this example also reveals that in Theorem 1.1(d)(i) the '$<$'-sign is possible.

If $(x_n)_{n\in\mathbb{N}}$ is a non-decreasing sequence, the following recursive formula for F_n^k is useful for practical computations, namely

$$\begin{aligned} F_n^k(x_1,\dots,x_{n-k}) &= 1-\sum_{j=0}^{n-k-1} P(X_{1:n}\le x_1,\dots,X_{j:n}\le x_j, X_{j+1:n}>x_{j+1}) \\ &= 1-\sum_{j=0}^{n-k-1}\binom{n}{j} F_j(x_1,\dots,x_j)\alpha_{j+1}^{n-j}, \end{aligned}$$

where $F_j \equiv F_j^0$ with $F_0^0\equiv 1$.

The proof of Theorem 1.1 is presented in Section 2. In the first part we summarize some useful preparatory analytic and probabilistic results allowing an immediate conclusion of the assertions of Theorem 1.1, except for part (b), which is proved by a different method. Some counterexamples, applications and concluding remarks in Section 3 finish the paper.

2. Proof of the main result

First notice that attention can be restricted to the special case of uniformly distributed random variables. For, if $U_1, \ldots, U_n$ are independent, identically distributed random variables having a uniform distribution on $[0,1]$, and the joint distribution function of $U_{1:n}, \ldots, U_{n-k:n}$ is denoted by G_n^k, then we have (cf. Reiss (1989), p. 17)

$$F_n^k(x_1, \ldots, x_{n-k}) = G_n^k(F(x_1), \ldots, F(x_{n-k})) = G_n^k(1-\alpha_1, \ldots, 1-\alpha_{n-k}),$$

whence it suffices to consider the assertion of Theorem 1.1 in terms of G_n^k. In the sequel we use the notation

$$H_n^k(\alpha_1, \ldots, \alpha_{n-k}) = G_n^k(1-\alpha_1, \ldots, 1-\alpha_{n-k}) \quad \text{and} \quad H_n \equiv H_n^0.$$

Before starting with the main part of the proof of Theorem 1.1 we first state some useful facts. To set notation, let $0 < \theta < 1/2$, $j_n = [\theta n]$, $k_n = [(1-\theta)n]$, $n \in \mathbb{N}$, where $[x]$ denotes the largest integer less than or equal to $x \in \mathbb{R}$. Define $\gamma_n = \max\{\alpha_{k_n}, \ldots, \alpha_{n-k}\}$ and $\nu_n = \max\{\alpha_{j_n}, \ldots, \alpha_{k_n-1}\}$ for $n - k \geq k_n > j_n$. Then it is easy to establish that

$$\text{(2.1)} \qquad \lim_{n\to\infty} n\alpha_n = c \quad \text{implies} \quad \lim_{n\to\infty} n\gamma_n = c/(1-\theta),$$

$$\text{(2.2)} \qquad \limsup_{n\to\infty} n\alpha_n = c \quad \text{implies} \quad \limsup_{n\to\infty} n\gamma_n = c/(1-\theta),$$

$$\text{(2.3)} \qquad \liminf_{n\to\infty} n\alpha_n = c \quad \text{implies} \quad \liminf_{n\to\infty} n\gamma_n \geq c/(1-\theta),$$

and

(2.4) if $(\alpha_n)_{n\in\mathbb{N}}$ is non-increasing and $\liminf_{n\to\infty} n\alpha_n = c$, then $\liminf_{n\to\infty} n\gamma_n = c/(1-\theta)$.

Furthermore,

(2.5) if $i_n \in \mathbb{N}$, $d_n \in [0,1]$, $n \in \mathbb{N}$, such that $\lim_{n\to\infty} i_n/n = d \in (0,1)$ and $\lim_{n\to\infty} d_n = d_0 > d$, then $\lim_{n\to\infty} P(U_{i_n:n} > d_n) = 0$.

This follows from the fact that $U_{i_n:n}$ tends to d in probability as n approaches infinity. Now, define $\delta = \inf_{n\in\mathbb{N}}(1-\alpha_n)$. Then it obviously holds

$$\text{(2.6)} \qquad \begin{aligned} &P(U_{j_n-1:n} \leq \delta, U_{k_n-1:n} \leq 1-\nu_n, U_{n-k:n} \leq 1-\gamma_n) \\ &\leq H_n^k(\alpha_1, \ldots, \alpha_{n-k}) \\ &\leq P(U_{n-k:n} \leq 1-\alpha_{n-k}). \end{aligned}$$

Elementary set algebra reveals that the lower bound in (2.6) is not less than

$$P(U_{n-k:n} \leq 1-\gamma_n) - P(U_{j_n-1:n} > \delta) - P(U_{k_n-1:n} > 1-\nu_n).$$

If $\limsup_{n\to\infty} n\alpha_n < \infty$, then $\lim_{n\to\infty} \alpha_n = \lim_{n\to\infty} \nu_n = 0$ and $\delta > 0$, hence for all $\theta \in (0, \min\{1/2, \delta\})$ (2.5) yields

$$\lim_{n\to\infty} P(U_{j_n-1:n} > \delta) = 0 \quad \text{and} \quad \lim_{n\to\infty} P(U_{k_n-1:n} > 1-\nu_n) = 0,$$

thus

$$\begin{aligned}\liminf_{n\to\infty} P(U_{n-k:n} \le 1-\gamma_n) &\le \liminf_{n\to\infty} H_n^k(\alpha_1, \dots, \alpha_{n-k}) \\ &\le \liminf_{n\to\infty} P(U_{n-k:n} \le 1-\alpha_{n-k}),\end{aligned} \tag{2.7}$$

and

$$\begin{aligned}\limsup_{n\to\infty} P(U_{n-k:n} \le 1-\gamma_n) &\le \limsup_{n\to\infty} H_n^k(\alpha_1, \dots, \alpha_{n-k}) \\ &\le \limsup_{n\to\infty} P(U_{n-k:n} \le 1-\alpha_{n-k}).\end{aligned} \tag{2.8}$$

Armed with these facts, we can easily conclude the assertions of Theorem 1.1, except for part (b), where different arguments are used.

(a) For $\lim_{n\to\infty} n\alpha_n = c < \infty$, part (a) can be proved as follows. Let $\theta \in (0, \min\{1/2, \delta\})$, then $\lim_{n\to\infty} n\gamma_n = c/(1-\theta)$ by (2.1). Now, (1.3), (2.7) and (2.8) imply $g_k(c/(1-\theta)) \le \liminf_{n\to\infty} H_n^k(\alpha_1, \dots, \alpha_{n-k}) \le \limsup_{n\to\infty} H_n^k(\alpha_1, \dots, \alpha_{n-k}) \le g_k(c)$. Letting $\theta \to 0$ yields the assertion. If $c = \infty$, define $\beta_n = \min\{\alpha_n, K/n\}$. In view of $\beta_n \le \alpha_n$ and $\lim_{n\to\infty} n\beta_n = K$, the assertion follows from the case $c < \infty$ by letting $K \to \infty$.

(c) First let $c < \infty$ and $\limsup_{n\to\infty} n\alpha_n = c$. Then (2.2) implies $\limsup_{n\to\infty} n\gamma_n = c/(1-\theta)$. With (1.5) and (2.7) we then obtain $g_k(c/(1-\theta)) \le \liminf_{n\to\infty} H_n^k(\alpha_1, \dots, \alpha_{n-k}) \le g_k(c)$, thus $\liminf_{n\to\infty} H_n^k(\alpha_1, \dots, \alpha_{n-k}) = g_k(c)$ by letting $\theta \to 0$. The case $c = \infty$ can be treated as in (a). The reverse direction in (c) is due to the fact that g_k is bijective.

(d) (i) Follows immediately by applying (1.4) and the right-hand side inequality of (2.8), which is also valid for $c = \infty$.

(d) (ii) First let $c < \infty$ and $\liminf_{n\to\infty} n\alpha_n = c$. Then (2.4), (1.4) together with (2.8) yield $\limsup_{n\to\infty} H_n^k(\alpha_1, \dots, \alpha_{n-k}) \ge g_k(c)$ by letting $\theta \to 0$, thus $\limsup_{n\to\infty} H_n^k(\alpha_1, \dots, \alpha_{n-k}) = g_k(c)$ by (d)(i). The case $c = \infty$ is a special case of (a), and finally the equivalence follows as in (c).

(b) In view of (c) and (d)(ii) we can restrict attention to the case $c < \infty$. From (c) we obtain $c = \limsup_{n\to\infty} n\alpha_n$. Now, assume that $c_1 = \liminf_{n\to\infty} n\alpha_n < c \in (0, \infty)$. Let $\beta_n = \max\{\alpha_n, \min\{1/2, c/n\}\}$, $n \in \mathbb{N}$, then $\lim_{n\to\infty} n\beta_n = c$ and $\beta_n \ge \alpha_n$ for all $n \in \mathbb{N}$. Since $c_1 = \liminf_{n\to\infty} n\alpha_n$, we can extract a subsequence $(n_i)_{i\in\mathbb{N}}$ of $\mathbb{N}$ such that $\lim_{i\to\infty}(n_i - k)\alpha_{n_i-k} = c_1$ and $\alpha_{n_i-k} < \alpha_{n_i-k-1}$ for all $i \in \mathbb{N}$. Then we obtain

$$\begin{aligned}&H_{n_i}^k(\alpha_1, \dots, \alpha_{n_i-k}) \\ &\quad\ge H_{n_i}^k(\beta_1, \dots, \beta_{n_i-k-1}, \alpha_{n_i-k}) \\ &\quad= P(U_{1:n} \le 1-\beta_1, \dots, U_{n_i-k-1:n} \le 1-\beta_{n_i-k-1}, U_{n_i-k:n} \le 1-\beta_{n_i-k})\end{aligned}$$

$$
\begin{aligned}
&+ P(U_{1:n} \le 1-\beta_1, \ldots, U_{n_i-k-1:n} \le 1-\beta_{n_i-k-1},\\
&\qquad\qquad 1-\beta_{n_i-k} < U_{n_i-k:n} \le 1-\alpha_{n_i-k})\\
&= H^k_{n_i}(\beta_1, \ldots, \beta_{n_i-k-1}, \beta_{n_i-k})\\
&+ \binom{n_i}{k+1}(\beta^{k+1}_{n_i-k} - \alpha^{k+1}_{n_i-k})H_{n_i-k-1}(\beta_1, \ldots, \beta_{n_i-k-1}).
\end{aligned}
$$

Passing to the limit for $i \to \infty$ on both sides of this inequality together with part (a) of Theorem 1.1 yields

$$g_k(c) \ge g_k(c) + \frac{1}{(k+1)!}(c^{k+1} - c_1^{k+1})g_0(c) > g_k(c),$$

which obviously is a contradiction. Hence the proof is complete. □

3. Concluding remarks

As announced at the end of Section 1, we investigate part (b) of Theorem 1.1 for $c = \infty$ and part (d)(i) for non-monotonic sequences $(x_n)_{n\in\mathbb{N}}$ in detail.

First of all it should be mentioned that for $c = \infty$ (1.2) always implies $\limsup_{n\to\infty} n\alpha_n = c = \infty$ (cf. Theorem 1.1(c)). Now, let F be the cdf of a non-degenerate probability distribution and $x_0 \in \mathbb{R}$ with $0 < F(x_0) < 1$. Define $\alpha_n = 1 - F(x_0)$ for $n \in \mathbb{N}$ odd and $\alpha_n = 0$ for $n \in \mathbb{N}$ even, then (1.2) is satisfied with $c = \infty$, but $\liminf_{n\to\infty} n\alpha_n = 0$, which means that (1.1) is not valid. Furthermore, if F is continuous, for every $d \in [0,\infty)$ it is possible to find a sequence $(x_n)_{n\in\mathbb{N}}$ or $(\alpha_n)_{n\in\mathbb{N}}$ such that (1.2) holds with $c = \infty$, but $\liminf_{n\to\infty} n\alpha_n = d$; consider e.g. $\alpha_n = \min\{1/2, d/n\}$ for $n \in \mathbb{N}$ even, and $\alpha_n = 1/2$ for $n \in \mathbb{N}$ odd. Moreover, $\limsup_{n\to\infty} H^k_n(\alpha_1, \ldots, \alpha_{n-k}) = 0 < g_k(d)$, which shows that in part (d)(i) the '$<$'-sign is possible.

A non-trivial example which illustrates (for continuous F) the existence of a sequence $(\alpha_n)_{n\in\mathbb{N}}$ with $g_k(\limsup_{n\to\infty} n\alpha_n) = \liminf_{n\to\infty} H^k_n(\alpha_1, \ldots, \alpha_{n-k}) < \limsup_{n\to\infty} H^k_n(\alpha_1, \ldots, \alpha_{n-k}) < g_k(\liminf_{n\to\infty} n\alpha_n)$ is given by $\alpha_n = \min\{1/2, c_1/n\}$ for $n \in \mathbb{N}$ even and $\alpha_n = \min\{1/2, c_2/n\}$ for $n \in \mathbb{N}$ odd, $0 < c_1 < c_2 < \infty$. Similarly as in the proof of part (b) it can be shown that $\limsup_{n\to\infty} H^k_n(\alpha_1, \ldots, \alpha_{n-k}) = g_k(c_2) + (c_2^{k+1} - c_1^{k+1})g_0(c_2)/(k+1)!$. A more detailed but analogous argumentation reveals that the last expression is not only $\le g_k(c_1)$, but in fact $< g_k(c_1)$. Since $g_k(c)$ can be interpreted as the probability $P_c(Z \le k)$, where Z has a Poisson distribution with parameter $c \in (0,\infty)$ under P_c, we obtain for all $k \in \mathbb{N}_0$, $0 < c_1 < c_2 < \infty$, the inequality

$$P_{c_1}(Z \le k) - P_{c_2}(Z \le k) > \frac{1}{(k+1)!}(c_2^{k+1} - c_1^{k+1})\exp(-c_2).$$

It is obvious from (1.3), (1.4) and (1.5) that the conditions $\lim_{n\to\infty} n\alpha_n = c$, $\liminf_{n\to\infty} n\alpha_n = c$ and $\limsup_{n\to\infty} n\alpha_n = c$ can be reformulated equivalently via conditions concerning the cdf of the $(k+1)$-th largest order statistic, namely $\lim_{n\to\infty} P(X_{n-k:n} \le x_{n-k}) = g_k(c)$, $\limsup_{n\to\infty} P(X_{n-k:n} \le x_{n-k}) = g_k(c)$ and

$\liminf_{n\to\infty} P(X_{n-k:n} \le x_{n-k}) = g_k(c)$. Hence, the joint distribution function of $X_{1:n}, \ldots, X_{n-k:n}$ from a sample $X_1, \ldots, X_n$ behaves just like the $(k+1)$-th largest order statistic $X_{n-k:n}$ if n is large (at least in the case where $(x_n)_{n\in\mathbb{N}}$ is a non-decreasing sequence, which is a quite natural assumption).

If F belongs to the domain of attraction of an extreme value distribution function Q, then the equivalence in (1.3) can easily be derived from the well-known equivalence

$$\lim_{n\to\infty} n(1 - F(a_n x + b_n)) = -\log Q(x) \qquad \text{for all} \qquad x \in \mathbb{R},$$

$$\text{iff} \qquad \lim_{n\to\infty} P(X_{n-k:n} \le a_n x + b_n) = Q(x) \sum_{j=0}^{k} (-\log Q(x))^j / j! \qquad \text{for all} \qquad x \in \mathbb{R}$$

for some normalizing constants $a_n > 0$ and $b_n \in \mathbb{R}$ (cf. Reiss ((1989), (5.1.4) and (5.1.28)), or Leadbetter *et al.* ((1983), Theorem 1.5.1 and Theorem 2.2.2)).

If $\lim_{n\to\infty} n\alpha_n$ exists, Theorem 1.1 may be applied to approximate expressions like $P(X_{i:n} \le x_i, i \in J_n^k)$, where $J_n^k \subseteq \{1, \ldots, n-k\}$ with $n-k \in J_n^k$. For, since $\{X_{1:n} \le x_1, \ldots, X_{n-k} \le x_{n-k}\} \subseteq \{X_{i:n} \le x_i, i \in J_n^k\} \subseteq \{X_{n-k:n} \le x_{n-k}\}$, it is obvious that $\lim_{n\to\infty} P(X_{i:n} \le x_i, i \in J_n^k) = g_k(c)$.

Examples of sequences $(\alpha_n)_{n\in\mathbb{N}}$ for which $\lim_{n\to\infty} n\alpha_n$ exists and which play an important role in multiple comparisons are $\alpha_n = 1 - (1-\alpha_n)^{1/n}$ with $\lim_{n\to\infty} n\alpha_n = -\log(1-\alpha)$ and $\alpha_n' = \alpha/n$, $\alpha \in (0,1)$, $n \in \mathbb{N}$. In the first case we obtain e.g. for $k = 0$ $\lim_{n\to\infty} H_n(\alpha_1, \ldots, \alpha_n) = 1 - \alpha$, in the second case $\lim_{n\to\infty} H_n(\alpha_1', \ldots, \alpha_n') = \exp(-\alpha)$.

It may appear strange to the reader that we did not restrict attention to non-decreasing sequences $(x_n)_{n\in\mathbb{N}}$ (or to non-increasing sequences $(\alpha_n)_{n\in\mathbb{N}}$) in advance. One reason for this was the following monotonicity problem first raised in Dunnett and Tamhane (1992):

> If $\alpha \in (0,1)$ is fixed, does there exist a *decreasing* sequence $(\alpha_n)_{n\in\mathbb{N}}$ such that $H_n(\alpha_1, \ldots, \alpha_n) = 1 - \alpha$ for all $n \in \mathbb{N}$?

A discussion of this question and some applications of Theorem 1.1 in multiple comparisons can be found in Finner *et al.* (1993). Recently, the monotonicity problem described above was solved by Dalal and Mallows (1992), who encountered it in a completely different context, independently of Dunnett and Tamhane (1992). They proved that for every $\alpha \in (0,1)$ there indeed exists a *strictly decreasing* sequence $(\alpha_n)_{n\in\mathbb{N}}$ such that $H_n(\alpha_1, \ldots, \alpha_n) = 1 - \alpha$ for all $n \in \mathbb{N}$.

Acknowledgements

The authors would like to thank two referees for their detailed comments and helpful suggestions which led to a considerable improvement of the presentation of the paper.

References

Dalal, S. R. and Mallows, C. L. (1992). Buying with exact confidence, *The Annals of Applied Probability*, **2**, 752–765.

Dunnett, C. W. and Tamhane, A. C. (1992). A step-up multiple test procedure, *J. Amer. Statist. Assoc.*, **86**, 162–170.

Finner, H., Hayter, A. J. and Roters, M. (1993). On a conjecture concerning the joint distribution function of order statistics with special reference to stepwise multiple test procedures, Tech. Report, No. 93-19, FB IV Mathematik/Informatik, University of Trier, Germany.

Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*, Springer, New York.

Reiss, R. D. (1989). *Approximate Distributions of Order Statistics*, Springer, New York.