Abstract. We introduce the notion of the $q$-analog of the $k$-th order statistics. We give a distribution and asymptotic distributions of $q$-analogs of the $k$-th order statistics and the intermediate order statistics with $r \to \infty$ and $r - k \to \infty$ in the projective geometry $PG (r - 1, q)$.

2000 AMS Mathematics Subject Classification: Primary: 60G70; Secondary: 51E20.

Key words and phrases: Order statistics, limit theorems, $q$-analog, finite projective geometry.

1. INTRODUCTION

The main results of this paper are asymptotic distributions of the $q$-analogs of the $k$-th minimal order statistics (Theorem 2.1) and the intermediate order statistics (Theorem 2.2). Theorem 2.2 generalizes the Theorem from [5] and Fact 3 from [6]. This paper is an extension (with full proofs) of the results announced in [7].

Let $GF (q)$ be a Galois field, where $q$ is the power of prime. Let $V (r, q)$ be an $r$-dimensional vector space over $GF (q)$. There exists a one-to-one correspondence between $k$-dimensional subspaces of projective geometry $PG (r - 1, q)$ and $k$-dimensional subspaces of the space $V (r, q)$. “Directions” in $V (r, q)$ are points of the projective geometry $PG (r - 1, q)$ of dimension $r - 1$. The subspace of dimension $k - 1$ has the rank $k$. For example, a line has a dimension one, but it has a rank two. Let $\sigma (A)$ denote the subspace spanned by $A$, i.e. the smallest subspace including $A$. Let $\rho (A)$ denote the rank of $\sigma (A)$. The monograph by Hirschfeld [2] gives a detailed exposition of this subject; see also [9] or [11]. Let $q$ be fixed and $n$ be a nonnegative integer. We use the standard notation $[n] = (q^n - 1)/(q - 1)$ (see, for example, [3] or [4]). It is well known (see [2]) that the projective geometry $PG (r - 1, q)$ has $[r]$ elements.

Projective geometries can be defined in an axiomatic way. A projective geometry satisfies the following axioms:

(1) Any two distinct points are on exactly one line.
Let \( q \) results are the \( q \)-analogs of known ones in the theory of extremal order statistics. In the case when \( q \) is the power of the prime, the subspaces of rank \( k \) in \( PG(r-1, q) \) are \( q \)-analogs of \( k \)-element sets. In such a meaning, our results are the \( q \)-analogs of known ones in the theory of extremal order statistics.

Let a sequence of random variables \( X_1, X_2, \ldots, X_n \) be given. We define the order statistics \( Z_k^{(n)} \), \( k = 1, 2, \ldots, n \), as random variables which are functions of random vector \( (X_1, X_2, \ldots, X_n) \) defined as follows. For any event \( \omega \), we arrange a sequence of realizations \( X_1(\omega) = x_1, X_2(\omega) = x_2, \ldots, X_n(\omega) = x_n \) in a non-decreasing sequence \( z_1 \leq z_2 \leq \ldots \leq z_n \). In this sequence, \( z_k \) is the realization of the random variable \( Z_k^{(n)} \), i.e. \( Z_k^{(n)}(\omega) = z_k \). For a fixed integer \( k \), the random variables \( Z_k^{(n)} \) are the \( k \)-th minimal order statistics and the random variables \( Z_{n-k+1}^{(n)} \) are the \( k \)-th maximal order statistics.

For the case of the projective geometry \( PG(r-1, q) \) we shall take \( n = [r] \). Let \( \{X_e\} \) be independent, identically distributed random variables with distribution function \( F(x) \) and assigned to elements of \( PG(r-1, q) \). Let us order the elements \( e_1, e_2, \ldots, e_r \) of \( PG(r-1, q) \) so that \( e_i \) has weight \( Z_i \). Let \( (Y_1, Y_2, \ldots, Y_r) \) be a subsequence of the sequence \( (Z_1, Z_2, \ldots, Z_n) \) such that \( Y_i^{(n)} = Z_k^{(n)} \) (for simplicity of the notation we sometimes write \( Y_i \) or \( Z_k \)). Let \( k_i \) be the least index with \( e_{k_i} \notin \sigma(e_{k_1}, e_{k_2}, \ldots, e_{k_{i-1}}) \). Note that \( k_1 = 1, k_2 = 2 \), i.e. \( Y_1 = Z_1, Y_2 = Z_2 \) and \( k_i \geq i \) for \( i \geq 3 \). The random variables \( Y_1, Y_2, \ldots, Y_r \) will be called the \( q \)-analogs of the order statistics.

For the better clarity of further formulas, we consider \((k+1)\)-st order statistics, \( k = 0, 1, \ldots, \) instead of \( k \)-th one, \( k = 1, 2, \ldots \)

\[ \Pr(Y_{k+1}^{(n)} > x) = \sum_{m=k}^{[k]} \left( \prod_{l=k}^{m-1} \frac{[k] - l}{[r] - l} \right) \left( \frac{[r]}{[r] - m} \right) \sum_{t=0}^{m} \binom{n}{t} \left( F(x) \right)^t (1 - F(x))^{n-t} \]
Proof. Note that

\[(1.1)\quad p_1 = \frac{[k] - l}{[r] - l}\]

is a probability that a point belongs to the space spanned by \(l\) earlier points, because \([k]\) denotes the number of elements of rank-\(k\) space, \([r]\) is a number of all elements and \(l\) means the number of earlier chosen elements. Then

\[(1.2)\quad p_2 = \prod_{l=k}^{m-1} \frac{[k] - l}{[r] - l}\]

is a probability that successively chosen points belong to a space determined by earlier chosen points, so we have to choose a next point, and

\[(1.3)\quad p_3 = \frac{[r] - [k]}{[r] - m}\]

is a probability that an \(m\)-th point does not belong to a space determined by earlier points, i.e. it spans a space of higher dimension, and

\[(1.4)\quad p_{4,t} = \binom{n}{t} (F(x))^t (1 - F(x))^{n-t}\]

are probabilities that exactly \(t\) points have weights smaller than \(x\). Combining (1.1), (1.2), (1.3) and (1.4) we conclude that

\[p_1 p_2 p_3 \sum_{t=1}^{m} p_{4,t} = \left( \prod_{l=k}^{m-1} \frac{[k] - l}{[r] - l}\right) \frac{[r] - [k]}{[r] - m} \sum_{t=0}^{m} \binom{n}{t} (F(x))^t (1 - F(x))^{n-t}\]

is a probability that \(m\) is an index of a point with the smallest weight, which does not belong to rank-\(k\) space, spanned by earlier chosen points. ■

2. LIMIT DISTRIBUTIONS

In this section, using known results concerning simple order statistics and limit distributions of random subsets of finite projective spaces, we will find limit distribution of \(q\)-analogos of order statistics.

We standardize random variables \(Z_{k}^{(n)}\) as follows:

\[\tilde{Z}_{k}^{(n)} = \frac{Z_{k}^{(n)} - b_n}{a_n}\]

with constants \(a_n > 0, b_n\) appropriately chosen, \(k\) fixed, and \(n\) increasing infinitely, N. W. Smirnov (see, for example, [8]) has shown that nondegenerate asymptotic
distributions of the normalized $k$-th minimal order statistics $\tilde{Z}_k^{(n)}$ can be of three types only:

(2.1) \[ \Psi_1^{(k)}(x) = 1 - P(k, \exp(x)), \quad -\infty < x < \infty, \]

(2.2) \[ \Psi_2^{(k)}(x) = \begin{cases} 0, & x \leq 0, \\ 1 - P(k, x^\alpha), & x > 0, \alpha > 0, \end{cases} \]

(2.3) \[ \Psi_3^{(k)}(x) = \begin{cases} 1 - P(k, (-x)^{-\alpha}), & x < 0, \alpha > 0, \\ 1, & x > 0, \end{cases} \]

where

(2.4) \[ P(k, \lambda) = \sum_{j=0}^{k-1} \frac{\lambda^j}{j!} \exp(-\lambda), \quad \lambda > 0. \]

Now we investigate a limit behaviour of a $q$-analog of the $k$-th minimal order statistics.

**Theorem 2.1.** For independent random variables with a distribution $F(x)$ a distribution of a $q$-analog of the $k$-th order statistics when $n \to \infty$ is given by

(2.5) \[ \Pr\left( \frac{Y_{k+1}^{(n)} - b_n}{a_n} < x \right) \to \Psi_i(x), \quad i = 1, 2, 3, \]

where a function $\Psi$ is defined by formulas (2.1)–(2.3).

**Proof.** Replacing $x$ by $a_nx + b_n$ in Proposition 1.1 we get

(2.6) \[ \Pr\left( \frac{Y_{k+1}^{(n)} - b_n}{a_n} > x \right) = \sum_{m=k}^{[k]} \left( \prod_{l=k}^{m-1} \frac{[k] - l}{[r] - l} \right) \frac{[r] - [k]}{[r] - m} \times \sum_{t=0}^{m} \left( \begin{array}{c} n \\vdash t \\ t \end{array} \right) (F(a_nx + b_n))^t \left( 1 - F(a_nx + b_n) \right)^{n-t} \]

\[ = \sum_{m=k+1}^{[k]} \left( \prod_{l=k}^{m-1} \frac{[k] - l}{[r] - l} \right) \frac{[r] - [k]}{[r] - m} + \left( \prod_{l=k}^{m-1} \frac{[k] - l}{[r] - l} \right) \frac{[r] - [k]}{[r] - k} \times \sum_{t=0}^{m} \left( \begin{array}{c} n \\vdash t \\ t \end{array} \right) (F(a_nx + b_n))^t \left( 1 - F(a_nx + b_n) \right)^{n-t}. \]

Using an asymptotic distribution (see formulas (2.1)–(2.3)) and the fact that when $n = [r] \to \infty$

\[ \prod_{l=k}^{m-1} \frac{[k] - l}{[r] - l} \to 0, \quad \frac{[r] - [k]}{[r] - m} \to 1, \quad \prod_{l=k}^{m-1} \frac{[k] - l}{[r] - l} \to 1, \quad \frac{[r] - [k]}{[r] - k} \to 1, \]
q-analogs of order statistics

we get
\[
\Pr \left( \frac{Y_{k+1}^{(n)} - b_n}{a_n} < x \right) \rightarrow \Psi_i(x), \quad i = 1, 2, 3. \quad \blacksquare
\]

For fixed \( k \), as \( n = [r] \rightarrow \infty \) the asymptotic distribution of the \( q \)-analog of the \( k \)-th order statistics coincides with the distribution of the simple \( k \)-th order statistics. This is because the number \( n = [r] \) of points of the projective geometry \( PG(r-1,q) \) is exponentially growing in \( r \rightarrow \infty \) (\( q \) is fixed) so that, for \( i \ll r \), the points \( e_1, e_2, \ldots, e_i \) are such that each \( e_i \) is, with probability tending to one, independent of \( e_1, e_2, \ldots, e_{i-1} \). Thus, for \( k \) fixed, the \( k \)-th minimal order statistics \( Y_k \) is asymptotically equal to the \( k \)-th order statistics \( Z_k \).

It is also interesting to consider the cases when \( k = k_n \rightarrow \infty \) as \( n = [r] \rightarrow \infty \), which can be called the cases of increasing ranks (see [8]). Two particular rates of increase are of special interest:

1. \( k_n \rightarrow \infty \) and \( k_n/n \rightarrow 0 \), which is called the case of intermediate ranks (the intermediate order statistics);
2. \( k_n/n \sim \Theta (0 < \Theta < 1) \), which is called the case of central ranks (the central order statistics).

If \( \{k_n\} \) is a non-decreasing intermediate order statistics sequence and there are constants \( a_n > 0 \) and \( b_n \) such that \( \Pr (a_n(Z_n^{(k_n)} - b_n) \leq x) \rightarrow L(x) \) for a nondegenerate distribution \( L \), then \( L \) has one of the three forms:

\[
L_1(x) = \begin{cases} 
\Phi(-a \log(-x)), & x < 0, a > 0, \\
1, & x \geq 0,
\end{cases}
\]

\[
L_2(x) = \begin{cases} 
0, & x \leq 0, a > 0, \\
\Phi(a \log x), & x > 0,
\end{cases}
\]

\[
L_3(x) = \Phi(x), \quad -\infty < x < \infty,
\]

where \( \Phi (\tau) \) is a Gaussian distribution function with zero mean and variance one.

Define a discrete random process \( \omega_r (k) \) as a Markov chain of subsets of elements of the \( PG(r-1,q) \), which starts with empty set and for \( k = 1, 2, \ldots, n = [r] \), \( \omega_r (k) \) is obtained by addition to \( \omega_r (k-1) \) a new, randomly chosen element of \( PG(r-1,q) \). In [5] (see also [6]) Kordecki and Łuczak have shown that for \( n = [r] \) if \( k-r \rightarrow \infty \), then \( \rho(\omega_r(k)) = r \) almost surely, whereas for \( r-k \rightarrow \infty \) we have \( \rho(\omega_r(k)) = k \) almost surely, \( q \)-analogs of the intermediate order statistics and the central order statistics \( (k/r \rightarrow 0 \text{ or } k/r \rightarrow \theta, 0 < \theta < 1) \) are expressed by the intermediate (“normal”) order statistics, because then \( k/n \rightarrow 0 \) for \( n = [r] \).

Now we define \( q \)-analogs of order statistics when \( k \rightarrow \infty \). Let \( Y_k^{(n)} \), where \( n \rightarrow \infty, k \rightarrow \infty \), but \( k/n \rightarrow 0 \), be a \( q \)-analog of an intermediate order statistics. Let \( Y_k^{(n)} \), where \( k \rightarrow \infty, n \rightarrow \infty, k/n \sim \Theta (0 < \Theta < 1) \), be a \( q \)-analog of a central order statistics.
THEOREM 2.2. For independent random variables with a distribution $F(x)$, a distribution of a $q$-analog of an intermediate order statistics, where $r \to \infty$ when $n = [r] \to \infty$ and $k \to \infty$, is expressed as

$$\Pr \left( \frac{Y_{k+1}^{(n)} - b_n}{a_n} < x \right) \to L_i(x), \quad i = 1, 2, 3,$$

where the functions $L_i$ are defined by formulas (2.7)–(2.9).

Proof. Consider once again the equation (2.6) in the proof of Theorem 2.1. Because first factors of the products

$$\prod_{l=k}^{m-1} \frac{[k]-l}{[r]-l}$$

are the greatest and

$$0 < \prod_{l=k}^{m-1} \frac{[k]-l}{[r]-l} < 1,$$

we have

$$\prod_{l=k}^{m-1} \frac{[k]-l}{[r]-l} < \frac{[k]-k}{[r]-k}.$$

Similarly, because

$$0 < \frac{[r]-[k]}{[r]-m} < 1,$$

we get

$$0 < \sum_{m=k+1}^{[k]} \left( \prod_{l=k}^{m-1} \frac{[k]-l}{[r]-l} \right) \frac{[r]-[k]}{[r]-m}$$

$$= \left( \prod_{l=k}^{k} \frac{[k]-l}{[r]-l} \right) \frac{[r]-[k]}{[r]-k} \left( \prod_{l=k}^{k+1} \frac{[k+1]-l}{[r]-l} \right) \frac{[r]-[k]}{[r]-k} + \left( \prod_{l=k}^{k+1} \frac{[k+1]-l}{[r]-l} \right) \frac{[r]-[k]}{[r]-k}$$

$$+ \ldots + \left( \prod_{l=k}^{[k]-1} \frac{[k]-l}{[r]-l} \right) \frac{[r]-[k]}{[r]-k}$$

$$= \frac{[k]-k}{[r]-k} \frac{[r]-[k]}{[r]-k} \left( 1 + \frac{[k]-(k+1)}{[r]-(k+1)} + \frac{[k]-(k+1)}{[r]-(k+2)} + \frac{[k]-(k+2)}{[r]-(k+3)} \right)$$

$$+ \ldots + \frac{[k]-(k+1)}{[r]-(k+2)} \frac{[k]-(k+2)}{[r]-(k+3)} \ldots \frac{[k]-(k-1)}{[r]-[k]}.$$
Moreover,

\[
\sum_{m=k+1}^{[k]} \left( \prod_{l=k}^{m-1} \frac{[k] - l}{[r] - l} \right) \frac{[r] - [k]}{[r] - m} < \frac{[k] - k}{[r] - k} \frac{[r] - [k]}{[r] - (k + 1)} \left( 1 + \frac{[k] - k}{[r] - [k]} + \left( \frac{[k] - k}{[r] - [k]} \right)^2 + \ldots \right) + \left( \frac{[k] - k}{[r] - [k]} \right)^{[k] - k - 1} \frac{1 - (([k] - k)/([r] - [k]))^{[k] - k - 1}}{1 - ([k] - k)/([r] - [k])} \rightarrow 0
\]

when \( r - k \rightarrow \infty, r \rightarrow \infty, k \rightarrow \infty \), because

\[
\frac{[k] - k}{[r] - k} = \frac{(q^k - 1)/(q - 1) - k}{(q^r - 1)/(q - 1) - k} \approx \frac{q^k}{q^r} = q^{r-k} \rightarrow 0,
\]

and we obtain

\[
\frac{[r] - [k]}{[r] - (k + 1)} = \frac{1 - [k]/[r]}{1 - (k + 1)/[r]} \rightarrow 1
\]

because

\[
\frac{[k]}{[r]} = \frac{(q^k - 1)/(q - 1)}{(q^r - 1)/(q - 1)} \approx \frac{q^k}{q^r} = q^{r-k} \rightarrow 0,
\]

\[
\frac{k + 1}{[r]} = \frac{(q^r - 1)/(q - 1)}{q^r - 1} \approx \frac{k}{q^r-1} \rightarrow 0.
\]

When \( n = [r] \rightarrow \infty \) we have

\[
\frac{[r] - [k]}{[r] - m} \rightarrow 1, \quad \prod_{l=k}^{[k]-1} \frac{[k] - l}{[r] - l} \rightarrow 1, \quad \frac{[r] - [k]}{[r] - k} \rightarrow 1.
\]

Then, using an asymptotic distribution (see formulas (2.7)-(2.9)), we get

\[
\Pr \left( \frac{Y_{k+1}^{(n)} - b_n}{a_n} < x \right) \rightarrow L_i(x), \quad i = 1, 2, 3. \quad \blacksquare
\]

Note that from the assumption that \( r - k \rightarrow \infty, r \rightarrow \infty, k \rightarrow \infty \) we infer that

\[
\frac{k}{n} = \frac{k}{[r]} = \frac{k}{(q^r - 1)/(q - 1)} \approx \frac{k}{q^r-1} \rightarrow 0.
\]

By Theorem 2.2 we obtain another proof of Fact 3 from [6].

Theorem 2.2 solves a problem of an asymptotic distribution of a \( q \)-analogue of an intermediate order statistics when \( r - k \rightarrow \infty \). Problems of \( q \)-analogue of asymptotic distributions for central and maximal order statistics remain unsolved.
REFERENCES


University of Business in Wrocław
Wrocław, Poland
E-mail: wojciech.kordecki@handlowa.eu

Department of Operations Research
The Poznań University of Economics
Poznań, Poland
E-mail: anna.lyczkowska-hancki@ae.poznan.pl

Received on 24.5.2009;
revised version on 27.11.2009