CrossMark

An ergodic theorem for proportions of observations that fall into random sets determined by sample quantiles

Anna Dembińska[1]

Abstract Assume that a sequence of observations $(X_n; n \geq 1)$ forms a strictly stationary process with an arbitrary univariate cumulative distribution function. We investigate almost sure asymptotic behavior of proportions of observations in the sample that fall into a random region determined by a given Borel set and a sample quantile. We provide sufficient conditions under which these proportions converge almost surly and describe the law of the limiting random variable.

Keywords Near order statistic observations · Stationary processes · Quantiles · Conditional quantiles · Almost sure convergence

1 Introduction

Let $(X_n; n \geq 1)$ be a sequence of random variables (rv's) with common cumulative distribution function (cdf) F, and $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be the order statistics corresponding to the sample $(X_1, X_2, \ldots, X_n)$. For a Borel set $A \subset \mathbb{R}$ and $1 \leq k \leq n$ we define a rv

$$K_{k:n}(A) = \#\{j \in \{1, \ldots, n\}; \; X_{k:n} - X_j \in A\} \tag{1}$$

✉ Anna Dembińska
 dembinsk@mini.pw.edu.pl

[1] Faculty of Mathematics and Information Science, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warsaw, Poland

counting elements in the sample that fall into a random region determined by the set A and the order statistic $X_{k:n}$. In particular,

$$K_{n:n}(\{0\}) = \#\{j \in \{1, \ldots, n\}; \ X_j = X_{n:n}\}$$

counts ties for the maximum while the rv

$$K_{n:n}([0, a)) = \#\{j \in \{1, \ldots, n\}; \ X_j \in (X_{n:n} - a, X_{n:n}]\}, \ \text{where } a > 0,$$

is the number of so-called near maxima, that is the number of elements in the sample falling within the distance a of the current maximum. If $k = k_n$ changes with n in such a way that $k_n/n \to \lambda \in (0, 1)$, then the following two objects

$$K_{k_n:n}((-a, 0)) = \#\{j \in \{1, \ldots, n\}; \ X_j \in (X_{k_n:n}, X_{k_n:n} + a)\}$$

and

$$K_{k_n:n}((0, a)) = \#\{j \in \{1, \ldots, n\}; \ X_j \in (X_{k_n:n} - a, X_{k_n:n})\}$$

where $a > 0$, can be viewed as the numbers of observations registered into the open right and left neighborhoods of the sample λ-quantile, respectively.

Asymptotic properties of the rv $K_{k:n}(A)$ as $n \to \infty$ have been studied by many authors. Their investigations were often motivated by various applications of derived results. For example, Eisenberg et al. (1993), Brands et al. (1994), Qi (1997), Bruss and Grübel (2003), Eisenberg (2009) and Gouet et al. (2009) examined the limiting behavior of the rv $K_{n:n}(\{0\})$ under the assumption that $(X_n; n \geq 1)$ is a sequence of independent and identically distributed (iid) rv's concentrated on nonnegative integers. Their aim was to answer the question whether $\lim_{n \to \infty} \Pr(K_{n:n}(\{0\}) = 1)$ – the limit of probability of no ties for the maximum–exists and to solve other related problems. Since $K_{n:n}(\{0\})$ can be viewed as the number of winners in a game with n players whose scores are $X_1, X_2, \ldots, X_n$, their results give insight into the existence of the limit of probability that there is a single winner among the n players. Another field where the rv $K_{k:n}(A)$ was applied is actuarial mathematics. Li and Pakes (2001) and Hashorva (2003, 2004) described and studied a model of insurance claims in which $K_{n:n}(A)$ counts claims with sizes in prescribed distance from the current maximal insurance claim. Asymptotic properties of $K_{n:n}(A)$ are then useful to describe the long-term behavior of this model. Next, the rv $K_{k:n}(A)$ plays an important role in the theory of spacings—Pakes and Steutel (1997) and Dembińska et al. (2007) noted that the asymptotic properties of spacings can be deduced from these of $K_{k:n}(A)$. Yet another application of $K_{k:n}(A)$ is to construct estimators of various quantities describing the cdf F. Results concerning limiting behavior of $K_{k:n}(A)$ can be exploited to show desirable limiting properties of these estimators; see, for example, Hashorva (2003, 2004), Müller (2003), Hashorva and Hüsler (2004) and Iliopoulos et al. (2012).

When studying the asymptotic behavior of the rv $K_{k:n}(A)$ we obtain different results according as (1) k or $n - k$ is held fixed (extreme case) or (2) $k_n/n \to \lambda \in (0, 1)$ (central or quantile case). Most of the work in the literature is devoted to the extreme

case; see, for example, a short description of the developments in this area given in
Dembińska (2014b). Yet, over the period of the last ten years, several papers dealing
with the quantile case have appeared. Dembińska et al. (2007), Pakes (2009), Iliopoulos
et al. (2012), (Dembińska 2012a, b, 2014b), Hashorva et al. (2013) and Nagaraja et al.
(2015) presented various limiting properties of $K_{k_n:n}(A)$ in the quantile case under
the condition that $(X_n; n \geq 1)$ is a sequence of iid rv's. Dembińska and Jasiński
(2016) quited the iid assumption and, under a weaker hypothesis that $(X_n; n \geq 1)$
is a strictly stationary and ergodic sequence, discussed the existence of the almost
sure limit of the proportions $K_{k_n:n}(A)/n$ as $n \to \infty$. In particular they gave sufficient
conditions for $K_{k_n:n}(A)/n$ to converge or diverge almost surely and established the
limit in the case of convergence. The aim of this paper is to provide a generalization
of the convergence result to strictly stationary but not necessarily ergodic sequences.
Under the mild assumption that $(X_n; n \geq 1)$ forms a strictly stationary process, we will
give general conditions under which $K_{k_n:n}(A)/n$ converges almost surely as $n \to \infty$.
We will also describe the distribution of the limiting rv using the idea of conditional
quantile.

The paper is organized as follows. In Sect. 2, we start with some preliminaries on
conditional quantiles and ergodic theory. Next, in Sect. 3, under assumption that the
underlying sequence of rv's forms a strictly stationary process, we provide sufficient
conditions for the existence of the almost sure limit of the proportions $K_{k_n:n}(A)/n$ and
establish the law of the limiting rv. Finally in Sect. 4, we give examples of application
of theorems from Sect. 3 to some special cases of strictly stationary sequences. It is
worth pointing out that in our derivations we do not assume that the cdf F is continuous.
We present results which are also valid for discontinuous and discrete F.

Throughout the paper we focus on the central case so we require that $(k_n; n \geq 1)$
is a sequence of integers such that

$$1 \leq k_n \leq n \quad \text{for all } n \geq 1 \text{ and } k_n/n \to \lambda \in (0, 1) \text{ as } n \to \infty. \tag{2}$$

We assume that the rv's $X_n, n \geq 1$, exist in a probability space $(\Omega, \mathcal{F}, P)$. By $\mathbb{R}$ and $\mathbb{N}$
we denote the sets of real numbers and positive integers, respectively. We write $x - A$
for the set $\{x - a; a \in A\}$ and ∂A for the boundary of the set A. By $\xrightarrow{a.s.}$ we denote
almost sure convergence and $a.s.$ stands for almost surely. Moreover, when different
probability measures appear, to avoid confusion, we write $\xrightarrow{P\text{-a.s.}}$ and E_P for almost
sure convergence and expectation with respect to the measure P, respectively, and we
say that an event A is true $P - a.s.$ if $P(A) = 1$. Next, when confusion can arise, we
add a superscript to $K_{k_n:n}(a, b)$ so that $K_{k_n:n}^{\mathbb{W}}(a, b)$ indicates that this rv arises from
the sequence $\mathbb{W} = (W_n, n \geq 1)$. Finally, $I(\cdot)$ stands for the indicator function, that is
$I(x \in A) = 1$ if $x \in A$ and $I(x \in A) = 0$ otherwise.

2 Preliminaries

As mentioned in the Introduction, to express the distribution of the almost sure limit
of $K_{k_n:n}(A)/n$ we will use the concept of conditional quantiles.

Definition 1 Suppose X is a rv on a probability space $(\Omega, \mathcal{F}, P)$, $\mathcal{G} \subset \mathcal{F}$ is a σ-field and $\lambda \in (0, 1)$. We say that a rv Q_λ is a conditional λth quantile of X with respect to $\mathcal{G}$ and write $Q_\lambda = \pi_\lambda(X|\mathcal{G})$ if and only if the following two conditions are satisfied:

1. Q_λ is $\mathcal{G}$-measurable,
2. $P(X \geq Q_\lambda|\mathcal{G}) \geq 1 - \lambda$ and $P(X \leq Q_\lambda|\mathcal{G}) \geq \lambda$ $a.s.$

Here are some elementary properties of conditional quantiles, relevant to our applications.

Theorem 1 *Let X be a rv and $\mathcal{G} \subset \mathcal{F}$ be a σ-field. Then,*

1. *there exists at least one conditional λth quantile of X with respect to $\mathcal{G}$;*
2. *there exist conditional λth quantiles $\underline{\pi}_\lambda(X|\mathcal{G})$ and $\overline{\pi}_\lambda(X|\mathcal{G})$ of X with respect to $\mathcal{G}$ such that*

$$\underline{\pi}_\lambda(X|\mathcal{G}) \leq \pi_\lambda(X|\mathcal{G}) \leq \overline{\pi}_\lambda(X|\mathcal{G}) \; a.s.$$

   for every conditional λth quantile $\pi_\lambda(X|\mathcal{G})$ of X with respect to $\mathcal{G}$;
3. *if X is $\mathcal{G}$-measurable then for any conditional λth quantile $\pi_\lambda(X|\mathcal{G})$ of X with respect to $\mathcal{G}$ we have*

$$\pi_\lambda(X|\mathcal{G}) = X \; a.s.;$$

4. *if a version of $\pi_\lambda(X|\mathcal{G})$ is constant then it is equal to a (usual) λth quantile π_λ^X of X which is not necessarily unique.*

Parts 1–3 of Theorem 1 with $\lambda = 1/2$ were proved by Tomkins (1975); see his Theorems 1, 3 and 2(i), respectively. The same reasoning applies to the case of $\lambda \neq 1/2$. Part 4 of Theorem 1 is just an easy observation. For more properties of conditional quantiles, we refer the reader to Tomkins (Tomkins 1975, 1978) and Ghosh and Mukherjee (2006).

By part 2 of Theorem 1 we can introduce the following definition of uniqueness of conditional quantile.

Definition 2 We will say that a conditional λth quantile Q_λ of X given $\mathcal{G}$ is unique if, given any other conditional λth quantile $Q_\lambda^\star$ of X given $\mathcal{G}$, we have $Q_\lambda = Q_\lambda^\star$ $a.s.$, in other words if $\underline{\pi}_\lambda(X|\mathcal{G}) = \overline{\pi}_\lambda(X|\mathcal{G})$ $a.s.$

Let us point out two simple examples of unique conditional quantiles.

Example 1 If X is $\mathcal{G}$-measurable then by part 3 of Theorem 1 for any $\lambda \in (0, 1)$, a conditional λth quantile of X given $\mathcal{G}$ is unique.

Example 2 If $\mathcal{G}$ is the trivial σ-algebra $\{\emptyset, \Omega\}$, then for every rv X and $\lambda \in (0, 1)$, $\pi_\lambda(X|\mathcal{G})$ is constant and therefore by part 4 of Theorem 1 a conditional λth quantile of X given $\mathcal{G}$ is unique if and only if a (usual) λth quantile π_λ^X of X is unique, that is if and only if

$$\inf\{x \in \mathbb{R} : F(x) \geq \lambda\} = \sup\{x \in \mathbb{R} : F(x) \leq \lambda\},$$

where F is the cdf of X.

To state and prove our results we will also use some concepts and facts from the ergodic theory. Throughout this paper $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}), \mathbb{Q})$ denotes a probability triple, where $\mathbb{R}^{\mathbb{N}}$ is the space of sequences of real numbers $(x_1, x_2, \ldots)$, $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$ stands for the Borel σ-field of subsets of $\mathbb{R}^{\mathbb{N}}$ and $\mathbb{Q}$ is a stationary probability measure on $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$. A set $B \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$ is called almost invariant for $\mathbb{Q}$ if

$$\mathbb{Q}\left((B \backslash T^{-1}B) \cup (T^{-1}B \backslash B)\right) = 0,$$

where $T^{-1}B := \{(x_1, x_2, \ldots) \in \mathbb{R}^{\mathbb{N}} : (x_2, x_3, \ldots) \in B\}$. The class of all almost invariant events for $\mathbb{Q}$ is denoted by $\mathcal{I}$. We will use the following well-known properties of $\mathcal{I}$; see, for example, Durrett (2010, Chapter 6) and Shiryaev (1996, Chapter V).

Theorem 2 1. $\mathcal{I}$ is a σ-field.
2. A rv X on $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}), \mathbb{Q})$ is $\mathcal{I}$-measurable if and only if

$$X((x_1, x_2, \ldots)) = X((x_2, x_3, \ldots)) \text{ for } \mathbb{Q}\text{-almost every } (x_1, x_2, \ldots) \in \mathbb{R}^{\mathbb{N}}.$$

3. The measure $\mathbb{Q}$ is ergodic if and only if every $\mathcal{I}$-measurable rv is $\mathbb{Q}$-a.s. constant.

We conclude this section with a theorem describing the almost sure behavior of central order statistics from strictly stationary processes. This result will be used in the next section where it will enable us to establish conditions that guarantee the almost sure convergence of the proportions $K_{k_n:n}(A)/n$ as $n \to \infty$.

Theorem 3 Let Y be a rv on a probability space $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}), \mathbb{Q})$, where the probability measure $\mathbb{Q}$ is stationary. Suppose that the sequence of rv's $(Y_n, n \geq 1)$ is defined by

$$Y_i((x_1, x_2, \ldots)) = Y((x_i, x_{i+1}, \ldots)) \text{ for } (x_1, x_2, \ldots) \in \mathbb{R}^{\mathbb{N}} \text{ and } i \geq 1. \quad (3)$$

If $(k_n, n \geq 1)$ is a sequence of integers satisfying (2) and the conditional λth quantile $\pi_\lambda(Y|\mathcal{I})$ of Y given $\mathcal{I}$ is unique, then

$$Y_{k_n:n} \xrightarrow{\mathbb{Q}\text{-a.s.}} \pi_\lambda(Y|\mathcal{I}).$$

The proof of Theorem 3 was given by Dembińska (2014a).

3 The ergodic theorem for relative frequencies

If a sequence of integers $(k_n, n \geq 1)$ is such that (2) holds, then

$$\frac{K_{k_n:n}(A)}{n} \xrightarrow{\text{a.s.}} \Pr(X_1 \in \pi_\lambda - A) \text{ as } n \to \infty, \quad (4)$$

whenever $X_n, n \geq 1$, are iid rv's with unique λth quantile and the Borel set A satisfies $\Pr(X_1 \in \gamma_\lambda - \partial A) = 0$; see Dembińska (2012b). Moreover Dembińska and Jasiński (2016) proved that the above result remains true if we replace the iid assumption with

the assumption that $(X_n, n \geq 1)$ is a strictly stationary and ergodic process. The aim of this section is to provide a complete generalization of this result by quiting the ergodicity assumption. The most important novelty of this generalization is that in the stationary case the limit in (4) need not to be a constant as it is in the stationary and ergodic case. Namely in the stationary case the limit can be a non-degenerate rv.

To simplify the proof of the main result, we first limit our attention to sets A of the form $A = (a, b)$, where $-\infty \leq a < b \leq \infty$. To shorten notation, we will write $K_{k:n}(a, b)$ instead of $K_{k:n}((a, b))$.

Theorem 4 *Under the assumptions of Theorem* 3,

$$K_{k_n:n}^{\mathbb{Y}}(a, b)/n \xrightarrow{\mathbb{Q}\text{-a.s.}} \mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in (a, b)|\mathcal{I}\big) \quad \text{as } n \to \infty, \tag{5}$$

provided that

$$\mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in \partial(a, b)|\mathcal{I}\big) = 0. \tag{6}$$

Proof By the definition

$$\frac{K_{k_n:n}^{\mathbb{Y}}(a, b)}{n} = \frac{\sum_{i=1}^n I\left(Y_{k_n:n} - b < Y_i < Y_{k_n:n} - a\right)}{n}. \tag{7}$$

The idea is to use the classic strong ergodic theorem to establish the a.s. limit of the averages in the RHS of (7). Yet, we cannot apply this theorem directly, because the elements of the sum in (7) might not form a strictly stationary process. In order to solve this problem we will replace $Y_{k_n:n}$ by a suitably chosen $\mathcal{I}$-measurable rv.

Fix $\varepsilon > 0$. Since by Theorem 3,

$$Y_{k_n:n} \xrightarrow{\mathbb{Q}\text{-a.s.}} \pi_\lambda(Y|\mathcal{I}),$$

we have $\mathbb{Q}$-almost surely, for all sufficiently large n,

$$\frac{K_{k_n:n}^{\mathbb{Y}}(a, b)}{n} \leq \frac{\sum_{i=1}^n I\big(\pi_\lambda(Y|\mathcal{I}) - \varepsilon - b < Y_i < \pi_\lambda(Y|\mathcal{I}) + \varepsilon - a\big)}{n}$$

$$= \frac{1}{n} \sum_{i=1}^n I\big(-\varepsilon - b < Y_i - \pi_\lambda(Y|\mathcal{I}) < \varepsilon - a\big). \tag{8}$$

Let $Z = I\big(-\varepsilon - b < Y - \pi_\lambda(Y|\mathcal{I}) < \varepsilon - a\big)$ and

$$Z_i((x_1, x_2, \ldots)) = Z((x_i, x_{i+1}, \ldots)) \text{ for } (x_1, x_2, \ldots) \in \mathbb{R}^{\mathbb{N}} \text{ and } i \geq 1. \tag{9}$$

Then

$$Z_i((x_1, x_2, \ldots)) = I\big(-\varepsilon - b < Y - \pi_\lambda(Y|\mathcal{I}) < \varepsilon - a\big)((x_i, x_{i+1}, \ldots))$$
$$= I\big(-\varepsilon - b < Y((x_i, x_{i+1}, \ldots)) - \pi_\lambda(Y|\mathcal{I})((x_i, x_{i+1}, \ldots)) < \varepsilon - a\big)$$

$$= I\big(-\varepsilon - b < Y_i((x_1, x_2, \ldots)) - \pi_\lambda(Y|\mathcal{I})((x_1, x_2, \ldots)) < \varepsilon - a\big) \ \mathbb{Q}\text{-}a.s.$$
$$= I\big(-\varepsilon - b < Y_i - \pi_\lambda(Y|\mathcal{I}) < \varepsilon - a\big)((x_1, x_2, \ldots)),$$

the next-to-last equality being a consequence of (3) and part 2 of Theorem 2. Therefore (8) can be rewritten as

$$\frac{K^{\mathbb{Y}}_{k_n:n}(a, b)}{n} \leq \frac{1}{n} \sum_{i=1}^{n} Z_i \qquad \mathbb{Q}\text{-}a.s.,$$

where Z_i, $i \geq 1$, satisfy (9) and hence form a strictly stationary process on $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}), \mathbb{Q})$. Since $E_{\mathbb{Q}}(|Z|) \leq 1$, the classic strong ergodic theorem (see, for example, Durrett 2010, p. 333) gives

$$\frac{1}{n} \sum_{i=1}^{n} Z_i \ \overset{\mathbb{Q}\text{-a.s.}}{\longrightarrow} \ E_{\mathbb{Q}}(Z|\mathcal{I}) = E_{\mathbb{Q}}\big(I(-\varepsilon - b < Y - \pi_\lambda(Y|\mathcal{I}) < \varepsilon - a)|\mathcal{I}\big).$$

It follows that, for all sufficiently large n,

$$\frac{K^{\mathbb{Y}}_{k_n:n}(a, b)}{n} \leq E_{\mathbb{Q}}(Z|\mathcal{I}) + \varepsilon = E_{\mathbb{Q}}\big(I(-\varepsilon - b < Y - \pi_\lambda(Y|\mathcal{I}) < \varepsilon - a)|\mathcal{I}\big) + \varepsilon \quad \mathbb{Q}\text{-}a.s.,$$

and hence that

$$\limsup_{n \to \infty} \frac{K^{\mathbb{Y}}_{k_n:n}(a, b)}{n} \leq E_{\mathbb{Q}}\big(I(-\varepsilon - b < Y - \pi_\lambda(Y|\mathcal{I}) < \varepsilon - a)|\mathcal{I}\big) + \varepsilon \qquad \mathbb{Q}\text{-}a.s. \quad (10)$$

In the same manner we can show that

$$\liminf_{n \to \infty} \frac{K^{\mathbb{Y}}_{k_n:n}(a, b)}{n} \geq E_{\mathbb{Q}}\big(I(\varepsilon - b < Y - \pi_\lambda(Y|\mathcal{I}) < -\varepsilon - a)|\mathcal{I}\big) - \varepsilon \qquad \mathbb{Q}\text{-}a.s. \quad (11)$$

By (10), (11) and the countability of $\mathbb{N}$ we see that $\mathbb{Q}$-almost surely, for all $m \geq 1$,

$$\limsup_{n \to \infty} \frac{K^{\mathbb{Y}}_{k_n:n}(a, b)}{n} \leq E_{\mathbb{Q}}\big(I(-\tfrac{1}{m} - b < Y - \pi_\lambda(Y|\mathcal{I}) < \tfrac{1}{m} - a)|\mathcal{I}\big) + \tfrac{1}{m}$$

and

$$\liminf_{n \to \infty} \frac{K^{\mathbb{Y}}_{k_n:n}(a, b)}{n} \geq E_{\mathbb{Q}}\big(I(\tfrac{1}{m} - b < Y - \pi_\lambda(Y|\mathcal{I}) < -\tfrac{1}{m} - a)|\mathcal{I}\big) - \tfrac{1}{m}.$$

Letting $m \to \infty$ and using the dominated convergence theorem for conditional expectations, we get

$$\limsup_{n \to \infty} \frac{K^{\mathbb{Y}}_{k_n:n}(a, b)}{n} \le E_{\mathbb{Q}}\big(I(-b \le Y - \pi_\lambda(Y|\mathcal{I}) \le -a)|\mathcal{I}\big)$$
$$= \mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in [a, b]|\mathcal{I}\big) \qquad \mathbb{Q}\text{-}a.s.$$

and

$$\liminf_{n \to \infty} \frac{K^{\mathbb{Y}}_{k_n:n}(a, b)}{n} \ge E_{\mathbb{Q}}\big(I(-b < Y - \pi_\lambda(Y|\mathcal{I}) < -a)|\mathcal{I}\big)$$
$$= \mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in (a, b)|\mathcal{I}\big) \qquad \mathbb{Q}\text{-}a.s.$$

Assumption (6) now shows that

$$\mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in (a, b)|\mathcal{I}\big) \le \liminf_{n \to \infty} \frac{K^{\mathbb{Y}}_{k_n:n}(a, b)}{n}$$
$$\le \limsup_{n \to \infty} \frac{K^{\mathbb{Y}}_{k_n:n}(a, b)}{n} \le \mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in (a, b)|\mathcal{I}\big) \qquad \mathbb{Q}\text{-}a.s.,$$

which clearly forces (5), and the proof is complete. □

We present below our main result.

Theorem 5 *Under the assumptions of Theorem* 3,

$$K^{\mathbb{Y}}_{k_n:n}(A)/n \xrightarrow{\mathbb{Q}\text{-}a.s.} \mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in A|\mathcal{I}\big) \quad as \ n \to \infty, \tag{12}$$

provided that A is a Borel subset of real numbers satisfying

$$\mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in \partial A|\mathcal{I}\big) = 0. \tag{13}$$

Proof We will apply similar arguments to that used in the proof of Theorem 2.1 of Dembińska and Jasiński (2016).

Let us first recall that any open subset of real numbers can be represented as a countable union of disjoint open intervals. Hence, in particular

$$Int\, A = \bigcup_{j=1}^{\infty} (a_j, b_j), \quad \text{for some } -\infty \le a_1 \le b_1 \le a_2 \le b_2 \ldots,$$

where it is understood that $(a, a) = \emptyset$. Therefore

$$\liminf_{n \to \infty} \frac{K_{k_n:n}^{\mathbb{Y}}(A)}{n} \geq \liminf_{n \to \infty} \frac{K_{k_n:n}^{\mathbb{Y}}(Int\,A)}{n} = \liminf_{n \to \infty} \frac{K_{k_n:n}^{\mathbb{Y}}\left(\bigcup_{j=1}^{\infty}(a_j, b_j)\right)}{n}$$

$$= \liminf_{n \to \infty} \frac{\sum_{j=1}^{\infty} K_{k_n:n}^{\mathbb{Y}}(a_j, b_j)}{n} \geq \sum_{j=1}^{\infty} \liminf_{n \to \infty} \frac{K_{k_n:n}^{\mathbb{Y}}(a_j, b_j)}{n}$$

$$= \sum_{j=1}^{\infty} \mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in (a_j, b_j)|\mathcal{I}\big) \quad \mathbb{Q}\text{-}a.s.,$$

the last equality beeing a consequence of Theorem 4. By linearity and the monotone convergence theorem for conditional expectations, we get

$$\sum_{j=1}^{\infty} \mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in (a_j, b_j)|\mathcal{I}\big) = \mathbb{Q}\left(\pi_\lambda(Y|\mathcal{I}) - Y \in \bigcup_{j=1}^{\infty}(a_j, b_j)|\mathcal{I}\right)$$

$$= \mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in Int\,A|\mathcal{I}\big) = \mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in A|\mathcal{I}\big) \quad \mathbb{Q}\text{-}a.s.,$$

where the last equality follows from (13).

We have thus proved that for any Borel set A satisfying (13), we have

$$\liminf_{n \to \infty} \frac{K_{k_n:n}^{\mathbb{Y}}(A)}{n} \geq \mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in A|\mathcal{I}\big) \quad \mathbb{Q}\text{-}a.s. \tag{14}$$

Note that (14) gives

$$\limsup_{n \to \infty} \frac{K_{k_n:n}^{\mathbb{Y}}(A)}{n} = \limsup_{n \to \infty} \frac{n - K_{k_n:n}^{\mathbb{Y}}(\mathbb{R}\backslash A)}{n} = 1 - \liminf_{n \to \infty} \frac{K_{k_n:n}^{\mathbb{Y}}(\mathbb{R}\backslash A)}{n}$$

$$\leq 1 - \mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in \mathbb{R}\backslash A|\mathcal{I}\big) = \mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in A|\mathcal{I}\big) \quad \mathbb{Q}\text{-}a.s., \tag{15}$$

because $\partial(\mathbb{R}\backslash A) = \partial A$ and therefore (13) holds also with A replaced by $\mathbb{R}\backslash A$.

Combining (14) with (15) establishes the desired convergence. $\square$

Although Theorem 5 is formulated in terms of sequences of rv's defined on the probability space $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}), \mathbb{Q})$, it can be used to derive a more general result for strictly stationary sequences existing in any probability space.

Theorem 6 *Let $\mathbb{X} = (X_n; n \geq 1)$ be a strictly stationary sequence and let $\mathbb{Q}$ be the stationary measure on $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$ defined by*

$$\mathbb{Q}(B) = \Pr(\mathbb{X} \in B) \quad for \quad B \in \mathcal{B}(\mathbb{R}^{\mathbb{N}}). \tag{16}$$

On $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}), \mathbb{Q})$ let us define the rv $Y : \mathbb{R}^{\mathbb{N}} \mapsto \mathbb{R}$ by

$$Y((x_1, x_2, \ldots)) = x_1 \, for \, (x_1, x_2, \ldots) \in \mathbb{R}^{\mathbb{N}}. \tag{17}$$

If $(k_n, n \geq 1)$ is a sequence satisfying (2) and the conditional λth quantile $\pi_\lambda(Y|\mathcal{I})$ of Y given $\mathcal{I}$ is unique, and (13) holds, then there exists a rv W such that

$$K_{k_n:n}^{\mathbb{X}}(A)/n \xrightarrow{\text{a.s.}} W \quad as\ n \to \infty.$$

Moreover, the joint distribution of the rv W and $(K_{k_n:n}^{\mathbb{X}}(A)/n, n \geq 1)$ is the same as the joint distribution of $\mathbb{Q}(\pi_\lambda(Y|\mathcal{I}) - Y \in A|\mathcal{I})$ and $(K_{k_n:n}^{\mathbb{Y}}(A)/n, n \geq 1)$, where $\mathbb{Y} = (Y_n, n \geq 1)$ is defined in (3).

Proof First, let us recall that the almost sure convergence of the sequence $(X_n, n \geq 1)$ to a rv X is entirely determined by the joint distribution of $(X, X_1, X_2, \ldots)$, because

$$X_n \xrightarrow{\text{a.s.}} X \quad \text{if and only if} \quad \lim_{n\to\infty} \Pr\left(\sup_{j\geq n} |X_j - X| > \varepsilon\right) = 0 \text{ for every } \varepsilon > 0.$$

Next, note that the sequences $(X_n; n \geq 1)$ and $(Y_n; n \geq 1)$ have the same distribution since, for all $B \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$,

$$\Pr((X_1, X_2, \ldots) \in B) = \mathbb{Q}(B) = \mathbb{Q}((Y_1, Y_2, \ldots) \in B).$$

Therefore the $\mathbb{Q}$-almost sure convergence of the sequence $K_{k_n:n}^{\mathbb{Y}}(A)/n$ to $\mathbb{Q}(\pi_\lambda(Y|\mathcal{I}) - Y \in A|\mathcal{I})$ entails the almost sure convergence of $K_{k_n:n}^{\mathbb{X}}(A)/n$ to a rv W such that the joint distribution of $\left(W, (K_{k_n:n}^{\mathbb{X}}(A)/n, n \geq 1)\right)$ is the same as the joint distribution of $\left(\mathbb{Q}(\pi_\lambda(Y|\mathcal{I}) - Y \in A|\mathcal{I}), (K_{k_n:n}^{\mathbb{Y}}(A)/n, n \geq 1)\right)$.   □

We conclude this section with an observation that under some additional condition, assumption (13) in Theorems 5 and 6 can be relaxed as shown in the following result.

Theorem 7 *Theorems 5 and 6 continue to hold if we additionally assume that, for all sufficiently large n,*

$$Y_{k_n:n} = \pi_\lambda(Y|\mathcal{I}) \qquad \mathbb{Q}\text{-a.s.} \tag{18}$$

and replace (13) by the condition that the set $A \subset \mathbb{R}$ can be represented as

$$A = B \cup C, \ where\ B \cap C = \emptyset, \tag{19}$$

or

$$A = B \backslash C, \ where\ C \subset B, \tag{20}$$

where the Borel sets B and C satisfy

$$\mathbb{Q}(\pi_\lambda(Y|\mathcal{I}) - Y \in \partial B|\mathcal{I}) = 0, \tag{21}$$

$$\mathbb{Q}(\pi_\lambda(Y|\mathcal{I}) - Y \in C|\mathcal{I}) = 0. \tag{22}$$

Proof First note that

$$\frac{K_{k_n:n}^{\mathbb{Y}}(C)}{n} = \frac{\sum_{i=1}^{n} I(Y_{k_n:n} - Y_i \in C)}{n} \xrightarrow{\mathbb{Q}\text{-a.s.}} 0,$$

because, by (18), the stationarity of the measure $\mathbb{Q}$ and (22), we have, for all sufficiently large n,

$$\mathbb{Q}(Y_{k_n:n} - Y_i \in C) = \mathbb{Q}(\pi_\lambda(Y|\mathcal{I}) - Y \in C) = E_{\mathbb{Q}}\big(\mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in C|\mathcal{I}\big)\big) = 0.$$

Therefore Theorem 5 and (22) give

$$\frac{K_{k_n:n}^{\mathbb{Y}}(A)}{n} = \frac{K_{k_n:n}^{\mathbb{Y}}(B)}{n} \pm \frac{K_{k_n:n}^{\mathbb{Y}}(C)}{n} \xrightarrow{\mathbb{Q}\text{-a.s.}} \mathbb{Q}(\pi_\lambda(Y|\mathcal{I}) - Y \in B|\mathcal{I})$$
$$= \mathbb{Q}(\pi_\lambda(Y|\mathcal{I}) - Y \in A|\mathcal{I}),$$

where $+$ and $-$ correspond to conditions (19) and (20), respectively. □

4 Examples

We will apply results of the previous section to two families of strictly stationary processes. The first family consists of strictly stationary and ergodic sequences of rv's. Note that in particular it contains all sequences of iid rv's (Grimmet and Stirzaker 2004, p. 399). The second family is the class of sequences of identical rv's that is the class of sequences of perfectly dependent variates.

4.1 Strictly stationary and ergodic sequences

Let $\mathbb{X} = (X_n, n \geq 1)$ be a strictly stationary and ergodic process. Then the measure $\mathbb{Q}$ defined in (16) is not only stationary but also ergodic. Consequently, by part 3 of Theorem 2, the rv $\pi_\lambda(Y|\mathcal{I})$, where Y is defined in (17), is constant $\mathbb{Q}$-a.s., and hence by part 4 of Theorem 1

$$\pi_\lambda(Y|\mathcal{I}) = \pi_\lambda^Y \qquad \mathbb{Q}\text{-}a.s..$$

It follows that, for any Borel set B,

$$\mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in B|\mathcal{I}\big) = \mathbb{Q}\big(\pi_\lambda^Y - Y \in B|\mathcal{I}\big) = E_{\mathbb{Q}}\big(I(Y \in \pi_\lambda^Y - B)|\mathcal{I}\big) \qquad \mathbb{Q}\text{-}a.s.$$

Next, since the measure $\mathbb{Q}$ is ergodic, the conditional expectation $E_{\mathbb{Q}}\big(I(Y \in \pi_\lambda^Y - B)|\mathcal{I}\big)$ is constant $\mathbb{Q}$-a.s. and equal to $E_{\mathbb{Q}}\big(I(Y \in \pi_\lambda^Y - B)\big)$; see Grimmet and Stirzaker (2004, p. 400). Therefore

$$\mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in B|\mathcal{I}\big) = \mathbb{Q}(Y \in \pi_\lambda^Y - B) \qquad \mathbb{Q}\text{-}a.s.$$

But $\mathbb{Q}(Y \in \pi_\lambda^Y - B) = \Pr(X_1 \in \pi_\lambda^X - B)$ since Y and X_1 have the same distribution by (17) and (16).

Thus in the strictly stationary and ergodic case we have, for any Borel set B,

$$\mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in B|\mathcal{I}\big) = \Pr(X_1 \in \pi_\lambda^X - B) \qquad \mathbb{Q}\text{-}a.s.$$

and then Theorem 6 reduces to

Corollary 1 *Let* $\mathbb{X} = (X_n, n \geq 1)$ *be a strictly stationary and ergodic process and* $(k_n, n \geq 1)$ *be a sequence of integers satisfying* (2). *If the λth quantile* π_λ^X *of* X_1 *is unique and* A *is a Borel subset of real numbers such that*

$$P(X_1 \in \pi_\lambda^X - \partial A) = 0, \tag{23}$$

then we have

$$\frac{K_{k_n:n}(A)}{n} \xrightarrow{\text{a.s.}} P(X_1 \in \pi_\lambda^X - A) \ as \ n \to \infty. \tag{24}$$

Thus we have recovered Theorem 2.1 of Dembińska and Jasiński (2016). Similarly, we can show that Theorem 2(b) of Dembińska (2012b) can be deduced from Theorem 7.

4.2 Sequences of identical variates

Let X be some rv and $X_n = X$ for all $n \geq 1$. Clearly $(X_n, n \geq 1)$ is strictly stationary. Moreover, assume that $(k_n, n \geq 1)$ is a sequence of integers satisfying (2). Then, for any $n \geq 1$, we have $X_{k_n:n} = X$ and

$$\frac{K_{k_n:n}^{\mathbb{X}}(A)}{n} = \frac{\sum_{i=1}^n I(X_{k_n:n} - X_i \in A)}{n} = \begin{cases} 1 & \text{if } 0 \in A \\ 0 & \text{if } 0 \notin A \end{cases}$$
$$\xrightarrow{n \to \infty} \begin{cases} 1 & \text{if } 0 \in A \\ 0 & \text{if } 0 \notin A \end{cases}, \tag{25}$$

where A is a Borel subset of real numbers. Below we will show that the application of Theorems 6 and 7 leads to the same conclusion.

To this end recall that in the case of identical variates we have $\mathcal{I} = \mathcal{B}(\mathbb{R}^{\mathbb{N}})$, see, for example Dembińska (2014a). Therefore, by part 3 of Theorem 1, the conditional λth quantile of Y given $\mathcal{I}$ is unique and

$$\pi_\lambda(Y|\mathcal{I}) = Y \ \mathbb{Q} - a.s.$$

It follows that, for any Borel subset B of real numbers,

$$\mathbb{Q}\big(\pi_\lambda(Y|\mathcal{I}) - Y \in B|\mathcal{I}\big) = \mathbb{Q}\big(0 \in B|\mathcal{I}\big) = I(0 \in B)$$

and we see that Theorem 6 implies (25) provided that $0 \notin \partial A$.

To handle the case when $0 \in \partial A$ we will use a specialized extension of Theorem 6, namely Theorem 7. Note that (18) holds. If $0 \notin A$ then $A = B \cup C$, where $B = \emptyset$ and $C = A$ satisfy (21) and (22). Hence the assumptions of Theorem 7 are fulfilled and using this result we obtain

$$\frac{K^{\mathbb{X}}_{k_n:n}(A)}{n} \xrightarrow{\text{a.s.}} 0, \tag{26}$$

which agrees with (25). If in turn $0 \in A$, then we write

$$\frac{K^{\mathbb{X}}_{k_n:n}(A)}{n} = \frac{K^{\mathbb{X}}_{k_n:n}(\mathbb{R})}{n} - \frac{K^{\mathbb{X}}_{k_n:n}(\mathbb{R}\setminus A)}{n}. \tag{27}$$

Since $0 \notin \mathbb{R}\setminus A$, (26) gives $K^{\mathbb{X}}_{k_n:n}(\mathbb{R}\setminus A)/n \xrightarrow{\text{a.s.}} 0$. By (1), $K^{\mathbb{X}}_{k_n:n}(\mathbb{R})/n = 1$ for all positive integers n. From (27) it follows that

$$\frac{K^{\mathbb{X}}_{k_n:n}(A)}{n} \xrightarrow{\text{a.s.}} 1 - 0 = 1. \tag{28}$$

Thus we arrived at (25).

References

Brands JJAM, Steutel FW, Wilms RJG (1994) On the number of maxima of a discrete sample. Stat Probab Lett 20:209–218

Bruss FT, Grübel R (2003) On the multiplicity of the maximum in a discrete random sample. Ann Appl Probab 13:1252–1263

Dembińska A (2012a) Asymptotic properties of numbers of observations in random regions determined by central order statistics. J Stat Plan Inference 142:516–528

Dembińska A (2012b) Limit theorems for proportions of observations falling into random regions determined by order statistics. Aust N Z J Stat 54(2):199–210

Dembińska A (2014a) Asymptotic behavior of central order statistics from stationary sequences. Stoch Process Appl 124:348–372

Dembińska A (2014b) Asymptotic normality of numbers of observations in random regions determined by order statistics. Statistics 48:508–523

Dembińska A, Jasiński K (2016) Asymptotic behavior of proportions of observations in random regions determined by central order statistics from stationary processes. Statistics. doi:10.1080/02331888.2016.1265966

Dembińska A, Stepanov A, Wesołowski J (2007) How many observations fall in a neighborhood of an order statistic? Comm Stat–Theory Methods 36:851–867

Durrett R (2010) Probability: Theory and Examples, 4th edn. Cambridge University Press, Cambridge

Eisenberg B (2009) The numbers of players tied for the record. Stat Probab Lett 79:283–288

Eisenberg B, Stengle G, Strang G (1993) The asymptotic probability of a tie for first place. Ann Appl Probab 3:731–745

Ghosh YN, Mukherjee B (2006) On probabilistic properties of conditional medians and quantiles. Stat Probab Lett 76:1775–1780

Gouet R, López FJ, Sanz G (2009) Limit laws for the cumulative number of ties for the maximum in a random sequence. J Stat Plan Inference 139:2988–3000

Grimmet GR, Stirzaker DR (2004) Probability and Random Processes. Oxford University Press, New York

Hashorva E (2003) On the number of near-maximum insurance claim under dependence. Insur Math Econ 32:37–49

Hashorva E (2004) Bivariate maximum insurance claim and related point processes. Stat Probab Lett 69:117–128

Hashorva E, Hüsler J (2004) Estimation of tails and related quantities using the number of near-extremes. Comm Stat–Theory Methods 34:337–349

Hashorva E, Macci C, Pacchiarotti B (2013) Large deviations for proportions of observations which fall in random sets determined by order statistics. Methodol Comput Appl Probab 15:875–896

Iliopoulos G, Dembińska A, Balakrishnan N (2012) Asymptotic properties of numbers of observations near sample quantiles. Statistics 46:85–97

Li Y, Pakes A (2001) On the number of near-maximum insurance claims. Insur Math Econ 28:309–323

Müller S (2003) Tail estimation based on numbers of near m-extremes. Methodol Comput Appl Probab 5:197–210

Nagaraja HN, Bharah K, Zhang F (2015) Spacings around an order statistic. Ann Inst Stat Math 67:515–540

Pakes A (2009) Numbers of observations near order statistics. Aust N Z J Stat 51:375–395

Pakes A, Steutel FW (1997) On the number of records near the maximum. Aust J Stat 39:179–193

Qi Y (1997) A note on the number of maxima in a discrete sample. Stat Probab Lett 33:373–377

Shiryaev AN (1996) Probability, 2nd edn. Springer, New York

Tomkins RJ (1975) On conditional medians. Ann Probab 3:375–379

Tomkins RJ (1978) Convergence properties of conditional medians. Can J Stat 6:169–177