Statistics and Probability Letters 109 (2016) 99–105

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

The rate of convergence in Hoeffding's theorem and some applications[☆]

I.S. Borisov

Sobolev Institute of Mathematics, Novosibirsk State University, Novosibirsk, 630090, Russia

ARTICLE INFO

Article history:
Received 3 August 2015
Received in revised form 25 October 2015
Accepted 6 November 2015
Available online 21 November 2015

Keywords:
Hoeffding's theorem
Order statistics
Empirical process
Quantile transform

ABSTRACT

We study the rate of convergence for the distributions of the expected values of all order statistics based on samples of dependent observations when the sample size n tends to infinity. As a consequence we study the asymptotics of the nth partial sum of the variances of all order statistics under consideration.

1. Introduction

Let $X_1, X_2, \dots$ be a sequence of identically distributed random variables with a common distribution function $F(t)$ and a finite expectation. In general, we do not suppose that the sequence $\{X_i\}$ consists of independent or stationarily connected observations. Consider the variational series $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ based on the sample $X_1, X_2, \dots, X_n$. The following assertion was proved by Hoeffding (1953).

Theorem 1. *Let $\{X_i\}$ be independent identically distributed random variables. Let $g(t)$ be a continuous function and let $h(t)$ be a nonnegative convex function satisfying the conditions $|g(t)| \le h(t)$ for all $t \in \mathbb{R}$ and $\mathbf{E}h(X_1) < \infty$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} g(\mathbf{E}X_{(j)}) = \mathbf{E}g(X_1). \tag{1}$$

Notice that if we consider the distribution assigning the weights $1/n$ to each of the points $\{\mathbf{E}X_{(i)};\ i \le n\}$ then the relation (1) taking place, in particular, for all bounded continuous functions g, implies the weak convergence of these distributions to the distribution of X_1 as $n \to \infty$.

In Borisov (2014) it was proved[1] that the limiting relation (1) is also valid for dependent observations if, as $n \to \infty$,

$$\frac{1}{n^2} \sum_{i,j=1}^{n} \mathbf{P}(\max\{X_i, X_j\} \le t) \to F^2(t) \tag{2}$$

☆ Supported by the Russian Foundation for Basic Research, grants 13–01–12415-OFI-M, 13–01–00511, and 14-01-00220.
E-mail address: sibam@math.nsc.ru.

1 In Borisov (2014) the relation (2) contains a misprint: The left-hand side of this relation contains the multiplier $1/n$ instead of $1/n^2$.

http://dx.doi.org/10.1016/j.spl.2015.11.004
0167-7152/

for every $t \in \mathbb{R}$. In the present paper we will prove a refinement of (1) under condition (2) and stronger moment restrictions than those in Theorem 1, in particular, for bounded observations. Moreover, we will study in detail the problem in both the cases of independent observations and of the integrand $g(t) = t^2$, i. e., we will study the limiting behavior of the sum of the variances of $X_{(j)}$, $j \leq n$. The results of such a kind are of interest, for example, for Regression Analysis when we study models with random regressors represented as variational series based on samples of dependent observations (for example, see Borisov, 2013).

From a statistical standpoint, in the present paper we study the proximity between the distribution of an averaging variational series and the corresponding unknown true distribution.

2. Main results

The goal of the paper is to study the rate of convergence to zero of the residual

$$\Delta_n(g, F) := \left| \frac{1}{n} \sum_{j=1}^{n} g(\mathbf{E}X_{(j)}) - \mathbf{E}g(X_1) \right|.$$

We will study the problem for the class of locally Lipschitz functions of power-type growth. In the sequel we will write that $g \in \mathrm{Lip}(B, \alpha)$ if

$$|g(x) - g(y)| \leq B(1 + (|x| \vee |y|)^{\alpha})|x - y|$$

for all $x, y \in \mathbb{R}$, with some constants $B > 0$ and $\alpha \geq 0$. Without loss of generality we suppose that $B = 1$ and $g(0) = 0$.

Let $F_n(t) := \frac{1}{n} \sum_{j=1}^{n} I\{X_j \leq t\}$ be the empirical distribution function based on the sample $X_1, \ldots, X_n$, where $I(\cdot)$ is the indicator of an event. Denote

$$\delta_n(t) := \mathbf{E}|F_n(t) - F(t)|.$$

The following elementary two-sided estimate is valid:

$$\mathbf{E}\left(F_n(t) - F(t)\right)^2 \leq \delta_n(t) \leq \left(\mathbf{E}\left(F_n(t) - F(t)\right)^2\right)^{1/2}. \tag{3}$$

Notice that the left inequality follows from the estimate $|F_n(t) - F(t)| \leq 1$. Since

$$\begin{aligned}\mathbf{E}\left(F_n(t) - F(t)\right)^2 &= \frac{1}{n^2} \sum_{i,j=1}^{n} \mathbf{E}I\{X_i \leq t\}I\{X_j \leq t\} - F^2(t) \\ &= \frac{1}{n^2} \sum_{i,j=1}^{n} \mathbf{P}(\max\{X_i, X_j\} \leq t) - F^2(t),\end{aligned} \tag{4}$$

from (3) it follows that condition (2) is necessary and sufficient for the limiting relation $\lim_{n\to\infty} \delta_n(t) = 0$ for all $t \in \mathbb{R}$.

Adduce a few simple estimates for $\delta_n(t)$ which directly follow from (3) and (4) for some popular types of dependence of the observations.

Case (i). Let the sequence $\{X_i\}$ consist of *pairwise negatively associated* random variables (this restriction is weaker than the well-known negative association) that are determined by the property (see Joag-Dev and Proschan, 1983)

$$\mathbf{P}(X_i \leq t, X_j \leq s) \leq F(t)F(s) \quad \text{for all } t, s \in \mathbb{R}, \text{ and } i \neq j.$$

If $t = s$ then from here we obtain

$$\mathbf{P}(\max\{X_i, X_j\} \leq t) \leq F^2(t) \quad \text{for all } i \neq j.$$

Hence, in the case under consideration, the relations (3) and (4) imply the estimate

$$\delta_n^2(t) \leq \frac{1}{n} F(t)(1 - F(t)). \tag{5}$$

Case (ii). Let the sequence $\{X_i\}$ (not necessarily stationary) satisfy the *strong mixing condition* (see Ibragimov and Linnik, 1971) with coefficient

$$\alpha(k) := \sup_{i} \sup_{A \in \sigma(\leqslant i), B \in \sigma(\geqslant i+k)} |\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)|,$$

where $\sigma(\leqslant i)$ and $\sigma(\geqslant i + k)$ are the σ-fields of the events generated by the families of random variables $\{X_j;\ j \leq i\}$ and $\{X_j;\ j \geq i + k\}$, respectively. Let, for all n,

$$\sum_{k=0}^{n} \alpha(k) \leq An^{\varepsilon}, \tag{6}$$