A BERRY-ESSÉEN BOUND FOR STUDENT'S STATISTIC IN THE NON-I.I.D. CASE

V. Bentkus[1] M. Bloznelis[1] F. Götze[1]

University of Bielefeld,
Vilnius University and the Institute of Mathematics and Informatics,
University of Bielefeld

November 1994

Abstract. We establish the Berry – Esséen bound for Student's statistic for independent nonidentically distributed random variables. If the third moments of these r.v. exist this bound yields an estimate which coincides (up to an absolute constant) with Esséen's bound for nonstudentized sums. If Lindeberg's condition for these r.v. holds, the bound yields the Central Limit Theorem for Student's statistics.

1. Introduction and Results

Let $X_1, ..., X_N$ be independent random variables. Let

$$\mathbf{t} = \overline{X}/\hat{\sigma}$$

denote the Student statistic, where

$$\overline{X} = N^{-1}(X_1 + \cdots + X_N), \qquad \hat{\sigma}^2 = N^{-1}\sum_{i=1}^{N}(X_i - \overline{X})^2.$$

[1]Research supported by the SFB 343 in Bielefeld.

1991 *Mathematics Subject Classification.* Primary 62E20; secondary 60F05.

Key words and phrases. Student's statistic, Berry-Esséen bound, nonidentically distributed random variables, convergence rate, Central Limit Theorem, self-normalized sums.

We shall investigate the rate of the normal approximation of $\sqrt{N}\mathbf{t}$ when the random variables $X_1, ..., X_N$ are nonidentically distributed. Define

$$\delta_N := \sup_x |\mathbf{P}(\sqrt{N}\ \mathbf{t} < x) - \Phi(x)|,$$

where $\Phi(x)$ is the standard normal distribution function. Denote

$$B_N := \sum_{i=1}^N \mathbf{Var} X_i \qquad \text{and} \qquad L_N := B_N^{-3/2} \sum_{i=1}^N \mathbf{E}|X_i|^3.$$

Theorem 1.1. *Let $X_1, ..., X_N$ denote independent random variables with mean zero. There exists an absolute constant $c > 0$ such that*

$$\delta_N \le c L_N, \tag{1.1}$$

whenever B_N is positive.

Inequality (1.1) is an extension to the case of studentized sums of the classical bound by Esséen (1945)

$$\sup_x |\mathbf{P}\left(X_1 + \cdots + X_N < x\sqrt{B_N}\right) - \Phi(x)| \le c L_N.$$

Theorem 1.1 is a consequence of a general result formulated as Theorem 1.2.

For a given number $V > 0$ define

$$Z_i = V^{-1/2} X_i \mathbb{I}\{X_i^2 \le V\}, \qquad 1 \le i \le N, \quad \text{and} \qquad M = \sum_{i=1}^N \mathbf{Var} Z_i, \tag{1.2}$$

where $\mathbb{I}\{A\}$ denotes the indicator function of event A.

Theorem 1.2. *Let $X_1, X_2, ..., X_N$ denote independent random variables. There exists an absolute constant $c > 0$ such that*

$$\begin{aligned} \delta_N \le c \sum_{i=1}^N \mathbf{P}(X_i^2 > V) &+ c M^{-1/2} |\sum_{i=1}^N \mathbf{E} Z_i| \\ &+ c M^{-1} \sum_{i=1}^N (\mathbf{E} Z_i)^2 + c M^{-3/2} \sum_{i=1}^N \mathbf{E}|Z_i - \mathbf{E} Z_i|^3, \end{aligned} \tag{1.3}$$

for any $V > 0$.

Theorem 1.2 is an extension of a Berry-Esséen bound for Student's statistic in the i.i.d. case due to Bentkus and Götze (1994) (see Corollary 1.2 below).