THE BERRY-ESSEEN BOUND FOR STUDENT'S STATISTIC

V. Bentkus[1] F. Götze[1]

University of Bielefeld

April 1994

Abstract. We prove the Berry-Esseen bound for the Student t-statistic. Under the assumption of a third moment this bound coincide (up to an absolute constant) with the classical Berry-Esseen bound for the mean. In general the distribution of the Student statistic converges to the standard normal distribution function at least as fast as the distribution of the mean, and sometimes faster. For example, rates of convergence can be proved if the underlying distribution is in the domain of attraction of the normal law.

Introduction and Results

Let $X_1, \dots, X_N$ denote i.i.d. mean zero random variables. Denote

$$\sigma^2 = \mathbf{E} X_1^2, \qquad \beta_s = \mathbf{E}|X_1|^s,$$

and

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

Assume that $\sigma > 0$ and define Student's statistic by

$$t = \bar{X}/\hat{\sigma}, \qquad \text{for} \quad N \geq 2,$$

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

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

Key words and phrases. Student's statistic, Berry-Esseen bound, convergence rate, Central Limit Theorem, self-normalized sums.

where $\bar{X}$ denotes the sample mean and $\hat{\sigma}$ the sample variance (if $\hat{\sigma} = 0$ then, for instance, put $t = 0$). The following Berry-Esseen bound is a consequence of our general estimates of the convergence rate of the distribution of $\sqrt{N}t$ to the standard normal distribution function. Denote

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

Theorem 1.1. *There exists an absolute constant $c > 0$ such that*

$$\delta_N \le \frac{c\beta_3}{\sigma^3\sqrt{N}}.$$

In the case the variance σ^2 is finite we have the following result.

Theorem 1.2. *There exists an absolute constant c such that*

$$\delta_N \le c\sigma^{-2}\mathbf{E}X_1^2\mathbf{I}\{X_1^2 > \sigma^2 N/2\} + cN^{-1/2}\sigma^{-3}\mathbf{E}|X_1|^3\mathbf{I}\{X_1^2 \le \sigma^2 N\}.$$

In particular, $\delta_N = o(N^{-s/2})$ provided $\beta_{2+s} < \infty$ and $0 \le s < 1$.

In order to formulate our main result we introduce a special norming and truncated moments. For a given random variable X_1 and a natural number N define the number $a^2 = a_N^2(X_1)$ by the truncated second moment equation

$$a^2 = \sup\{b : \ \mathbf{E}X_1^2\mathbf{I}\{X_1^2 \le bN\} \ge b\}, \qquad a \ge 0. \tag{1.1}$$

We shall need the following well-known properties of a.

Lemma 1.3. *The number a in (1.1) exists for any random variable X_1 and any N, and a is the largest solution of the equation*

$$a^2 = \mathbf{E}X_1^2\mathbf{I}\{X_1^2 \le a^2 N\}. \tag{1.2}$$

If $\sigma < \infty$ then $a \le \sigma$. If $\sigma = \infty$ then $a_N \to \infty$ as $N \to \infty$. If $\mathbf{P}\{X_1 = 0\} < 1$ then a is positive for sufficiently large N. If $\mathbf{E}X_1^2 = 1$ and $a^2 < 1/2$ then $\mathbf{E}X_1^2\mathbf{I}\{X_1^2 > N/2\} \ge 1/2$. Furthermore, $a^2(tX_1) = t^2a^2(X_1)$, for all $t \in \mathbf{R}$.

Denote

$$Y_i = a^{-1}N^{-1/2}X_i\mathbf{I}\{X_1^2 \le a^2 N\}, \qquad \text{for} \quad 1 \le i \le N.$$

Note that Y_i does not change if we replace X_i by τX_i, $\tau > 0$. Furthermore, $|Y_i| \le 1$ and $\mathbf{E}Y_i^2 = 1/N$.

Our main result is the following estimate which holds without assuming the existence of moments.