Math-Net.Ru

Общероссийский математический портал

М. Bloznelis, Н. Putter, Second-order and bootstrap approximation to Student's t -statistic, *Теория вероятн. и ее примен.*, 2002, том 47, выпуск 2, 374–381

DOI: <https://doi.org/10.4213/tvp3669>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.162.127.20

26 апреля 2023 г., 18:03:03

12. Фролов А. Н. О законе больших чисел Эрдеша–Реньи для неодинаково распределенных случайных величин при нарушении условия Крамера. — Вестник ЛГУ, 1991, в. 2 (№ 8), с. 57–61.
13. Фролов А. Н. Об асимптотическом поведении приращений сумм независимых случайных величин. — Вестник СПбГУ, 1993, в. 4 (№ 22), с. 45–48.
14. Петров В. В. Предельные теоремы для сумм независимых случайных величин. М.: Наука, 1987, 317 с.
15. Rozovsky L. V. On a lower bound for large deviation probabilities of sum of independent random variables with finite variances. — International Conference «Asymptotic Methods in Probability and Statistics»: Abstracts of communications. St. Petersburg, 1998, p. 236–237.

Поступила в редакцию
4.XI.1999

© 2002 г.

BLOZNELIS M.*, PUTTER H.**

SECOND ORDER AND BOOTSTRAP APPROXIMATION TO STUDENT'S t STATISTIC¹⁾

Справедливость для t -статистики Стьюдента разложения Эджворта до первого члена доказана при минимальных условиях: распределение наблюдений нерешетчато и имеет конечный третий момент. В качестве следствия получена корректность второго порядка метода бутстреп для t -статистики Стьюдента при этих оптимальных условиях, таким образом классический результат [19] распространяется на стьюдентизированное среднее.

Ключевые слова и фразы: критерий Стьюдента, разложение Эджворта, бутстреп, асимптотическое разложение, автонормированные суммы.

1. Introduction and results. Let $X_1, \dots, X_N, \dots$ be independent identically distributed random variables with the common distribution function F , expectation $\mathbf{E}X_1 = \mu$, and finite nonvanishing variance $\sigma^2 = \mathbf{E}(X_1 - \mu)^2$. Let

$$t_N = \frac{\bar{X}_N - \mu}{\hat{\sigma}_N}$$

denote Student's t statistic, where $\hat{\sigma}_N^2 = N^{-1} \sum_{i=1}^N (X_i - \bar{X}_N)^2$ and $\bar{X}_N = N^{-1}(X_1 + \dots + X_N)$. It is well known that the statistic $T_N = \sqrt{N}t_N$ is asymptotically standard normal, i.e., as $N \rightarrow \infty$

$$\sup_x |G_N(x) - \Phi(x)| \rightarrow 0, \quad \text{where } G_N(x) = \mathbf{P}\{T_N \leq x\} \quad (1.1)$$

denotes the distribution function of T_N and where $\Phi(x)$ denotes the standard normal distribution function.

The accuracy of (1.1) was studied by a number of authors, Helmers and van Zwet [16], Helmers [15], Slavova [20], Hall [13], Prašková [17], Friedrich [11], Bentkus and Götze [3], Bentkus, Bloznelis and Götze [2], etc.

* Vilnius University and the Institute of Mathematics and Informatics, Vilnius, Lithuania; e-mail: mindaugas.bloznelis@maf.vu.lt

** University of Leiden, The Netherlands; e-mail: H.Putter@amc.uva.nl

¹⁾ The first author supported by Lithuanian Science and Studies Foundation and A. von Humboldt Foundation; the second author supported by the SFB 343 in Bielefeld.

A Berry–Esseen bound of order $O(N^{-1/2})$ under the minimal moment condition $\mathbf{E}|X_1|^3 < \infty$ was obtained by Slavova [20]. Bentkus and Götze [3] proved the bound $\sup_x |G_N(x) - \Phi(x)| < cN^{-1/2}\beta_3/\sigma^3$, thus extending the classical result of Esseen [10] to the Studentized sums. Here $\beta_3 := \mathbf{E}|X_1 - \mu|^3$.

A higher order approximation to Student's t statistic was considered by Chung [9], Bhattacharya and Ghosh [5], Chibisov [8], Hall [12], Bentkus, Götze and van Zwet [4], Putter and van Zwet [18], etc.

Although increasingly general and precise, none of these results is optimal in the sense of expansions for $G_N(x)$ being established under minimal conditions. Hall [12] proved the validity of a k -term Edgeworth expansion for Student's t statistic with remainder $o(N^{-k/2})$, for every integer k , provided that $\mathbf{E}|X_1|^{k+2} < \infty$ and the distribution F is nonsingular. The moment conditions in Hall [12] are the minimal ones, but the smoothness condition on the distribution F of the observations is too restrictive. What is perhaps more important, Hall's result is only valid for a fixed underlying distribution function. As a result, it cannot be applied to the bootstrap.

The aim of the present paper is to prove the validity of an expansion under minimal conditions, in such a way that an extension to the bootstrap is straightforward. To this end, we approximate $G_N(x)$ by the one-term Edgeworth expansion

$$H_N(x) = \Phi(x) + \frac{\kappa_3}{6\sqrt{N}}(2x^2 + 1)\Phi'(x), \quad \kappa_3 = \frac{\mathbf{E}(X_1 - \mu)^3}{\sigma^3}$$

and construct explicit bounds for the remainder $\Delta_N = \sup_x |G_N(x) - H_N(x)|$.

The minimal smoothness condition which allows one to prove the validity of one-term Edgeworth expansion, i.e., to prove the bound $\Delta_N = o(N^{-1/2})$, is that F is nonlattice.

Theorem 1.1. *Assume that $\mathbf{E}|X_1|^3 < \infty$ and the distribution of X_1 is nonlattice. Then*

$$\Delta_N = o(N^{-1/2}) \quad \text{as } N \rightarrow \infty. \quad (1.2)$$

R e m a r k. If $\mu = 0$, then Theorem 1.1 remains valid if we replace Δ_N by $\tilde{\Delta}_N := \sup_x |\tilde{G}_N(x) - H_N(x)|$, where $\tilde{G}_N(x) := \mathbf{P}\{\tilde{T}_N \leq x\}$ denotes the distribution function of the selfnormalized sum

$$\tilde{T}_N = \frac{X_1 + \cdots + X_N}{(X_1^2 + \cdots + X_N^2)^{1/2}}.$$

B o o t s t r a p. Given $\mathbb{X}_N := \{X_1, \dots, X_N\}$, let $X_1^*, \dots, X_N^*$ denote independent random variables uniformly distributed in $\mathbb{X}$. Write

$$T_N^* = \sqrt{N} \frac{\overline{X_N^*} - \overline{X_N}}{\hat{\sigma}_N^*}, \quad (\hat{\sigma}_N^*)^2 = N^{-1} \sum_{i=1}^N (X_i^* - \overline{X_N^*})^2$$

and let $G_N^*(x) := \mathbf{P}\{T_N^* \leq x | \mathbb{X}_N\}$ denote the conditional probability of the event $\{T_N^* \leq x\}$ given $\mathbb{X}_N$, for $x \in \mathbf{R}$.

Theorem 1.2. *Suppose that $\mathbf{E}|X_1|^3 < \infty$ and that the distribution of X_1 is nonlattice. Then as $N \rightarrow \infty$*

$$\sup_{x \in \mathbf{R}} |G_N^*(x) - G_N(x)| = o(N^{-1/2}) \quad \text{a.s.} \quad (1.3)$$

Theorem 1.2 improves earlier results of Babu and Singh [1], Helmers [14] and Putter and van Zwet [18], where the bound (1.3) was established assuming that F is nonlattice and under increasingly sharp moment conditions, the sharpest to date being $\mathbf{E}|X_1|^{3+\varepsilon} < \infty$, for some $\varepsilon > 0$, obtained in the latter paper.

Theorems 1.1 and 1.2 are consequences of an explicit upper bound for Δ_N given in Proposition 2.1, which is formulated in Section 2 below.

A usual proof of the validity of Edgeworth expansions for distribution functions combines Esseen's [10] smoothing lemma and expansions of characteristic functions. Such scheme applies easily to a linear statistic since its characteristic function has a comparatively simple multiplicative structure. Before to apply it to Student's t statistic we approximate $\mathbf{t}$ by a statistic which is conditionally linear in the first m observations $X_1, \dots, X_m$,

$m < n$, given the remaining part of the sample $X_{m+1}, \dots, X_n$. Here $m \approx \ln N$. Bentkus and Götze [3] used similar idea to estimate the characteristic function of Student's t statistic.

The rest of the paper is organized as follows. In Section 2 we prove Theorems 1.1 and 1.2. Several steps of the proofs which are more technical are deferred to Section 3.

We are grateful to anonymous referee for remarks which improve the presentation.

2. Proofs. At the beginning of this section we formulate Proposition 2.1. Then we prove Theorems 1.1 and 1.2, which are corollaries of Proposition 2.1. The proof of our main result, Proposition 2.1, is postponed until after the proofs of Theorems 1.1 and 1.2.

Write

$$\rho_u = 1 - \sup \left\{ |\mathbf{E} \exp\{itX_1\}| : \frac{\sigma^2}{9\beta_3} \leq |t| \leq \frac{u}{\sigma} \right\}, \quad u > 0.$$

For the remainder of this paper we shall assume without loss of generality that $\mathbf{E}X_1 = 0$ and $\sigma^2 = \mathbf{E}X_1^2 = 1$.

Proposition 2.1. *There exists an absolute constant $c > 0$ such that for each $N = 2, 3, \dots$ and each $1 < u < N^{1/6}$,*

$$\Delta_N \leq \frac{c}{\sqrt{N}} \left(\frac{1}{u} + \frac{A_N}{\rho_u^4} + B_N \right), \quad (2.1)$$

$$A_N = N^{-1/2}(1 + \mathbf{E}X_1^4 \mathbb{I}\{X_1^2 \leq N\}), \quad B_N = N^{-1/2}\beta_3^2 + \mathbf{E}|X_1|^3 \mathbb{I}\{X_1^2 \geq N\}.$$

Proof of Theorem 1.1. It is easy to show that $\mathbf{E}|X_1|^3 < \infty$ implies $A_N = o(1)$ and $B_N = o(1)$ as $N \rightarrow \infty$. Furthermore, by the nonlattice property of F , we have $\rho_u > 0$, for every $u > 1$. Hence, one can find an increasing sequence $u_N \rightarrow \infty$ as $N \rightarrow \infty$ such that $\rho_{u_N}^4 A_N = o(1)$ and, therefore, the right-hand side of (2.1) with $u = u_N$ is of order $o(N^{-1/2})$.

Proof of Theorem 1.2. We shall apply (2.1) to $G_N^*(x)$ conditionally, given $\mathbb{X}_N$. Denote $Y_1 = (X_1^* - \bar{X}_N)/\hat{\sigma}_N$ and write $\hat{\kappa}_3 = \mathbf{E}^*Y_1^3$, $\hat{\beta}_3 = \mathbf{E}^*|Y_1|^3$, where $\mathbf{E}^*$ denotes the conditional expectation given $\mathbb{X}_N$.

Define

$$H_N^*(x) = \Phi(x) + \frac{\hat{\kappa}_3}{6\sqrt{N}} (2x^2 + 1) \Phi'(x).$$

Then by (2.1), we have for every $1 < u < N^{1/6}$,

$$\Delta_N^* := \sup_x |G_N^*(x) - H_N^*(x)| \leq \frac{c}{\sqrt{N}} \left(\frac{1}{u} + \frac{A_N^*}{\rho_u^4} + B_N^* \right), \quad (2.2)$$

$$A_N^* = N^{-1/2} \left(1 + \mathbf{E}^*Y_1^4 \mathbb{I}\{Y_1^2 \leq N\} \right),$$

$$B_N^* = N^{-1/2} \hat{\beta}_3^2 + \mathbf{E}^*|Y_1|^3 \mathbb{I}\{Y_1^2 \geq N\},$$

where we denote

$$\hat{\rho}_u = 1 - \sup \left\{ |\mathbf{E}^* \exp\{itY_1\}| : b_N \leq |t| \leq c_N \right\}, \quad b_N = \frac{\hat{\sigma}_N^2}{9\hat{\beta}_3}, \quad c_N = \frac{u}{\hat{\sigma}_N}.$$

In order to prove the theorem it suffices to show

$$\Delta_N^* = o(N^{-1/2}) \quad \text{a.s. as } N \rightarrow \infty. \quad (2.3)$$

Indeed, by the strong law of large numbers (SLLN), $\hat{\kappa}_3 \rightarrow \kappa_3$ a.s. as $N \rightarrow \infty$.

Hence, $\sup_x |H_N(x) - H_N^*(x)| = o(N^{-1/2})$ a.s. and this in combination with (1.2) and (2.3) implies (1.3).

Let us show (2.3). Given $a > 1$ write

$$\beta_3(a) = \mathbf{E}|X_1|^3 \mathbb{I}\{|X_1| \geq a\}, \quad \hat{\beta}_3(a) = \mathbf{E}^*|Y_1|^3 \mathbb{I}\{|Y_1| \geq a\},$$

and note that

$$\begin{aligned} A_N^* &\leq \frac{1}{\sqrt{N}} + \frac{a}{\sqrt{N}} \mathbf{E}^* |Y_1|^3 \mathbb{I}\{|Y_1| \leq a\} + \widehat{\beta}_3(a) \leq 2 \frac{a}{\sqrt{N}} \widehat{\beta}_3 + \widehat{\beta}_3(a), \\ B_N^* &\leq N^{-1/2} \widehat{\beta}_3^2 + \widehat{\beta}_3(a), \end{aligned} \quad (2.4)$$

where in the last inequality it is assumed that N is sufficiently large, namely $N \geq a^2$. Since $\mathbf{E}|X_1|^3 < \infty$, the SLLN implies that as $N \rightarrow \infty$

$$\widehat{\beta}_3 \rightarrow \beta_3 \quad \text{and} \quad \limsup_N \widehat{\beta}_3(a) \leq \beta_3(a) \quad \text{a.s.}, \quad (2.5)$$

$$b_N \rightarrow b := \frac{\sigma^2}{9\beta_3} \quad \text{and} \quad c_N \rightarrow \frac{u}{\sigma} \quad \text{a.s.} \quad (2.6)$$

Furthermore, Singh [19] showed that for any, but fixed $d > 0$,

$$\sup_{0 \leq |t| \leq d} |\mathbf{E} \exp\{itX_1\} - \mathbf{E}^* \exp\{itX_1^*\}| = o(1) \quad \text{a.s. as } N \rightarrow \infty,$$

see formula (2.4) therein. This fact in combination with the SLLN for $\widehat{\sigma}_N$, b_N and c_N , see (2.6), implies

$$\widehat{\rho}_u \rightarrow \rho_u \quad \text{a.s. as } N \rightarrow \infty, \quad (2.7)$$

for every fixed $u > 0$. Substitution of (2.4), (2.7) in (2.2) in combination with (2.5) gives

$$\limsup_N \sqrt{N} \Delta_N^* \leq c \left(u^{-1} + \beta_3(a)(1 + \rho_u^{-4}) \right) \quad \text{a.s.},$$

for arbitrary, but fixed $u > 1$ and $a > 0$. Since $\mathbf{E}|X_1|^3 < \infty$, we have $\beta_3(a) = o(1)$ as $a \rightarrow \infty$ and therefore the right-hand side can be made arbitrary small by choosing u and a sufficiently large. We obtain (2.3) thus completing the proof of the theorem.

Proof of Proposition 2.1. For clarity we start by outlining the main steps of the proof. We first use Lemma 3.1 below to replace the statistic T_N by a statistic S_0 , which is conditionally linear in the first m observations $X_1, \dots, X_m$, given the remaining observations of the sample, $X_{m+1}, \dots, X_N$. With $K_0(x) = \mathbf{P}\{S_0 \leq x\}$ denoting the distribution function of S_0 , an application of Berry–Esseen's smoothing lemma then reduces the problem of bounding $|K_0(x) - H_N(x)|$ to that of bounding the difference $|\widehat{K}_0(t) - \widehat{H}_N(t)|$, where $\widehat{K}_0$ and $\widehat{H}_N$ denote the Fourier transforms of K_0 and H_N respectively. The (conditional) linearity of S_0 produces a multiplicative component in $\widehat{K}_0$ and in combination with the smoothness condition (nonlattice condition on F) guarantees an exponential decay of $|\widehat{K}_0(t)|$, for large t , $|t| \geq c(F)\sqrt{N}$. Finally we bound the difference $|\widehat{K}_0(t) - \widehat{H}_N(t)|$, for $|t| \leq c(F)\sqrt{N}$.

Here and in what follows we write $c(a, b, \dots)$ to denote a constant that depends only on the quantities $a, b, \dots$. By $c, c_1, c_2, \dots$ we denote generic absolute constants. The expression $\exp\{x\}$ is abbreviated by $e\{x\}$. We shall write $A \ll B$ to denote the fact that $A \leq cB$. If Q denotes the sum $q_1 + \dots + q_k$ and $A \subset \{1, \dots, k\}$ then write $Q_A = \sum_{j \in A} q_j$. Given $A = \{i_1, \dots, i_m\} \subset \{1, \dots, N\}$ we write $\mathbf{E}_{i_1, \dots, i_m}$ to denote the conditional expectation given $\{X_j, j \notin A\}$. For a function h write $\|h\| = \sup_x |h(x)|$.

Let $g: \mathbf{R} \rightarrow \mathbf{R}$ be a function which is infinitely many times differentiable with bounded derivatives and such that

$$\frac{8}{9} \leq g(x) \leq \frac{8}{7}, \quad \text{for all } x \in \mathbf{R}, \quad \text{and} \quad g(x) = \frac{1}{\sqrt{x}}, \quad \text{for } \frac{7}{8} \leq x \leq \frac{9}{8}.$$

Write $c_g = \|g\| + \|g'\| + \|g''\| + \|g'''\|$.

First, we replace the random variables $X_1, \dots, X_N$ by truncated random variables $Y_i = a^{-1} N^{-1/2} X_i \mathbb{I}\{X_i^2 \leq a^2 N\}$, $i = 1, \dots, N$, where a is the largest nonnegative solution of the equation

$$a^2 = \mathbf{E} X_1^2 \mathbb{I}\{X_1^2 \leq a^2 N\}.$$

Note that $|Y_i| \leq 1$ and $\mathbf{E}Y_i^2 = 1/N$. Write $Y = Y_1 + \dots + Y_N$, $\eta = \eta_1 + \dots + \eta_N$, $\eta_i = Y_i^2 - \mathbf{E}Y_i^2$, $b_k = \mathbf{E}|Y_1|^k$, $\mathcal{M} = Nb_4$, $\gamma = N|\mathbf{E}Y_1|$, $\gamma_0 = N^{-1/2}\mathbf{E}|X_1|^3\mathbb{I}\{X_1^2 \geq N/2\}$. By Hölder's inequality,

$$b_3 \geq b_2^{3/2} = N^{-3/2}, \quad \mathcal{M} \geq Nb_2^2 = N^{-1}, \quad (b_3N)^2 \leq N^2b_2b_4 = \mathcal{M}. \quad (2.8)$$

We may and shall assume that for a sufficiently small $c_0 > 0$,

$$\frac{\beta_3}{\sqrt{N}} \leq c_0, \quad \rho_u^{-4}N^{-1} \leq c_0, \quad \gamma_0 \leq c_0. \quad (2.9)$$

Indeed, if the first inequality fails, the bound (2.1) follows from the simple inequality $\Delta_N \ll 1 + \beta_3/\sqrt{N} \ll \beta_3^2/N$. Hence, without loss of generality we may assume that $\beta_3/\sqrt{N} \leq c_0$. Then $\Delta_N \ll 1 + \beta_3/\sqrt{N} \ll 1$ and each of the remaining inequalities in (2.9) implies (2.1).

Using (2.9) and Lemma 3.2 we obtain $\frac{3}{4} < a \leq 1$ and, hence, $\mathcal{M} \leq A_N$. Therefore, in order to prove the proposition it suffices to show that

$$\Delta_N \ll \mathcal{R} + T^{-1}, \quad \text{where} \quad \mathcal{R} = \frac{\mathcal{M}}{\rho_u^4} + \gamma_0, \quad T = u\sqrt{N}. \quad (2.10)$$

Furthermore, (2.9) implies $\Delta_N \ll 1$ and therefore we may and shall assume without loss of generality that $\mathcal{M} \leq c_0$, since otherwise (2.10) follows from $\Delta_N \ll 1 \ll \mathcal{M}$.

Let m be the smallest integer such that $m \geq 4\rho_u^{-1} \ln N$. Write $A = \{1, \dots, m\}$ and $B = \{m+1, \dots, N\}$ and denote $V_B = 1 + \eta_B - Y_B^2/N$. Introduce the statistics

$$S = Yg(1 + \eta - Y^2/N), \quad S_0 = Yg(V_B) + \eta_A Y_B g'(V_B). \quad (2.11)$$

By Lemma 3.1, the probability $\mathbf{P}\{|S_0 - T_N| \geq N^{-2/3}\}$ is not greater than $c\mathcal{M}/\rho_u^2 + c\gamma_0$. Now Slutsky's argument gives

$$\Delta_N \leq \|K_0 - H_N\| + N^{-2/3}\|H'_N\| + c\mathcal{R}.$$

But $\|H'_N\| \leq c$, by (2.9). Hence, in order to prove (2.1) it remains to show $\|K_0 - H_N\| \ll \mathcal{R} + T^{-1}$.

We are going to apply Esseen's [10] smoothing lemma. We have

$$\|K_0 - H_N\| \ll I + T^{-1}, \quad \text{where} \quad I = \int_{0 \leq |t| \leq T} \frac{|\widehat{K}_0(t) - \widehat{H}_N(t)|}{|t|} dt.$$

Write $I = I_1 + I_2 + I_3$, where

$$I_1 = \int_{0 \leq |t| \leq L} \frac{|\widehat{K}_0(t) - \mathbf{E}e\{itS\}|}{|t|} dt, \quad I_2 = \int_{0 \leq |t| \leq L} \frac{|\mathbf{E}e\{itS\} - \widehat{H}_N(t)|}{|t|} dt, \\ I_3 = \int_{L \leq |t| \leq T} \frac{|\widehat{K}_0(t) - \widehat{H}_N(t)|}{|t|} dt, \quad L^{-1} = 4Nb_3.$$

We shall bound I_1 , I_2 and I_3 . An application of Lemma 3.3 gives $I_2 \ll \mathcal{R}$.

Let us show $I_1 \ll \rho_u^{-2}\mathcal{M} + N^{-2/3}$. By (3.2) and (3.3), see below, $|\widehat{K}_0(t) - \mathbf{E}e\{itS\}| \leq |t|\mathbf{E}|S_0 - S| \ll |t|\mathbf{E}R$, $R = R_1 + \dots + R_5$, where the random variables R_i are given by (3.2) and (3.3). A simple calculation shows $\mathbf{E}R \ll N^{-1/2}(\rho_u^{-2}\mathcal{M} + N^{-2/3})$. We obtain

$$I_1 \ll L\mathbf{E}R \ll \sqrt{N}\mathbf{E}R \ll \rho_u^{-2}\mathcal{M} + N^{-2/3}.$$

It remains to show that $I_3 \ll \mathcal{R}$. We have $I_3 \leq I_4 + I_5$,

$$I_4 = \int_{L \leq |t| \leq T} \frac{|\widehat{K}_0(t)|}{|t|} dt, \quad I_5 = \int_{L \leq |t| \leq T} \frac{|\widehat{H}_N(t)|}{|t|} dt.$$

By (2.9), $|\widehat{H}_N(t)| \leq e\{-t^2/2\}(1 + |t|^3)$. Therefore, $I_5 \leq e\{-(cL)^2\}$. The inequality $e\{-s^2\} \leq s^{-2}$ (which holds for sufficiently large s) applied to $s = cL$ gives $I_5 \ll (Nb_3)^2 \ll \mathcal{M}$, see (2.8). Finally, we shall show

$$I_4 \ll N^{-1}. \quad (2.12)$$

Write $\mathbb{I} = \mathbb{I}\{|Y_B| \leq N^{1/20}\}$. We have $|\widehat{K}_0(t)| \leq \mathbf{E}|\psi(t)| + \mathbf{P}\{|Y_B| > N^{1/20}\}$, $\psi(t) = \mathbf{E}_A e\{itS_0\}\mathbb{I}$. By Chebyshev's inequality, the probability is bounded by $\mathbf{E}|Y_B|^{40}/N^2 \ll N^{-2}$ (here we use the bound $\mathbf{E}|Y_B|^k \leq c(k)$, see, e.g., [3]). Furthermore,

$$|\psi(t)| \leq |\kappa|^m, \quad \kappa = \mathbf{E}_1 e\{it(Y_1 g(V_B) + \eta_1 Y_B g'(V_B))\}\mathbb{I}.$$

We complete the proof of (2.12) by showing

$$|\kappa| \leq 1 - \frac{\rho_u}{2}. \quad (2.13)$$

Indeed, this inequality implies $|\kappa|^m \leq e\{-m\rho_u/2\} \leq N^{-2}$, by the choice of m . Let us prove (2.13). Expanding in powers of $it\eta_1 Y_B g'(V_B)$ we obtain

$$\kappa = \mathbf{E}_1 e\{itY_1 g(V_B)\}\mathbb{I} + R,$$

$$|R| \leq 2T\mathbf{E}_1 |\eta_1 Y_B| c_g \mathbb{I} \leq 2c_g T N^{1/20-1} \leq 2c_g N^{-1/4} < \frac{\rho_u}{4}. \quad (2.14)$$

In the last step we used (2.9). Furthermore, write $Z = N^{-1/2} a^{-1} X_1 g(V_B)$. We have

$$\begin{aligned} |\mathbf{E}_1 e\{itY_1 g(V_B)\} - \mathbf{E}_1 e\{itZ\}| &\leq 2T\mathbf{E} |Y_1 g(V_B) - Z| \\ &\leq 2T\|g\| N^{-3/2} a^{-3} \mathbf{E}|X_1|^3 \mathbb{I}\{X_1^2 \geq a^2 N\} \leq 6\gamma_0 N^{-1/4} \leq N^{-1/4} \leq \frac{\rho_u}{4}. \end{aligned} \quad (2.15)$$

Here we used inequalities $T \leq N^{3/4}$, $2a^{-3}\|g\| \leq 6$ and $\gamma_0 \leq c_0$, see (2.9). But $\mathbf{E}_1 e\{itZ\} \leq 1 - \rho_u$. Collecting (2.14) and (2.15) we obtain (2.13) thus completing the proof of the proposition.

3. Appendix.

Lemma 3.1. Assume that (2.9) holds. Let S_0 be given by (2.11). Then

$$\mathbf{P}\{|S_0 - T_N| \geq N^{-2/3}\} \ll \mathcal{R}_1, \quad \mathcal{R}_1 = \rho_u^{-2} \mathcal{M} + \gamma_0. \quad (3.1)$$

P r o o f. The inequalities

$$\begin{aligned} \mathbf{P}\left\{\sqrt{N} \mathbf{t}_N \neq \frac{Y}{\sqrt{1 + \eta - Y^2/N}}\right\} &\leq N \mathbf{P}\{X_1^2 > a^2 N\} \leq \left(\frac{4}{3}\right)^3 \gamma_0, \\ \mathbf{P}\left\{\frac{Y}{\sqrt{1 + \eta - Y^2/N}} \neq S\right\} &\leq \mathbf{P}\left\{|\eta| > \frac{1}{4}\right\} + \mathbf{P}\left\{\frac{Y^2}{N} > \frac{1}{4}\right\} \end{aligned}$$

and

$$\mathbf{P}\left\{|\eta| > \frac{1}{4}\right\} \leq 16\eta^2 \leq 32N\mathbf{E}Y_1^4, \quad \mathbf{P}\left\{\frac{Y^2}{N} > \frac{1}{4}\right\} \leq 4N^{-1}\mathbf{E}Y^2 = \frac{4}{N}$$

imply $\mathbf{P}\{T_N \neq S\} \ll \mathcal{R}_1$.

In order to prove (3.1) it remains to show $\mathbf{P}\{|S - S_0| > N^{-2/3}\} \ll \mathcal{R}_1$. Write $S = Yg(V_B + \eta_A + W)$, where $W = -Y_A^2/N - 2Y_A Y_B/N$. Expanding g in powers of W we get $|S - Yg(V_B + \eta_A)| \leq c_g |YW|$, where $|YW| \leq R_1 + R_2 + R_3$ and

$$R_1 = \frac{|Y_A|^3}{N}, \quad R_2 = \frac{3Y_A^2|Y_B|}{N}, \quad R_3 = \frac{2|Y_A|Y_B^2}{N}. \quad (3.2)$$

Furthermore, expanding g in powers of η_A we obtain

$$|Yg(V_B + \eta_A) - S_0| \leq c_g(R_4 + R_5), \quad R_4 = |Y_A \eta_A|, \quad R_5 = |Y_B \eta_A^2|. \quad (3.3)$$

Then $|S - S_0| \leq c_g(R_1 + \dots + R_5)$ and, by Chebyshev's inequality,

$$\mathbf{P}\{|S - S_0| \geq N^{-2/3}\} \ll N^{2/3} \mathbf{E}R_1 + N^{2/3} \mathbf{E}R_2 + \mathbf{E}N^{4/3} R_3^2 + \mathbf{E}(N^{2/3} R_4)^{4/3} + N^{2/3} \mathbf{E}R_5.$$

A simple calculation shows that the right-hand side is bounded by $c\mathcal{R}_1$ thus completing the proof of (3.1).

Lemma 3.2. Assume that $\gamma_0 \leq \frac{1}{4}$. Then

$$1 - \sqrt{2}\gamma_0 \leq a \leq 1, \quad (3.4)$$

In particular we have $\frac{3}{4} \leq a \leq 1$.

P r o o f. Clearly, $a^2 \leq \sigma^2 = 1$. For $u \geq 0$, write $\phi(u) = \mathbf{E}X_1^2 \mathbb{I}\{X_1^2 \leq uN\}$. We have

$$\tau := 1 - \phi\left(\frac{1}{2}\right) = \mathbf{E}X_1^2 \mathbb{I}\left\{X_1^2 > \frac{N}{2}\right\} \leq \sqrt{2}\gamma_0 \leq \frac{1}{2}. \quad (3.5)$$

The function $\phi(u)$ is nondecreasing. Therefore, (3.5) implies $\phi(1 - \tau) \geq \phi(\frac{1}{2}) = 1 - \tau$. But $\phi(1) \leq \sigma^2 = 1$. Then there exists a solution of the equation $u = \phi(u)$ in the interval $1 - \tau \leq u \leq 1$. This implies $a^2 \geq 1 - \tau$ and we obtain $a \geq a^2 \geq 1 - \tau$. This inequality in combination with (3.5) yields (3.4) thus completing the proof of the lemma.

Lemma 3.3. Assume that (2.9) holds and that $\mathcal{M} \leq c_0$, where c_0 is a sufficiently small absolute constant. Then

$$I_2 = \int_{0 \leq |t| \leq L} \frac{|\mathbf{E}e\{itS\} - H_N(t)|}{|t|} dt \ll \mathcal{M} + \gamma_0, \quad L = (4b_3N)^{-1}.$$

P r o o f. Write $R = \mathbf{E}e\{itS\} - H_N(t)$. Split $I_2 = I_{2,1} + I_{2,2}$, where $I_{2,1} = \int_{0 \leq |t| \leq c} |t|^{-1} |R| dt$, $I_{2,2} = \int_{c \leq |t| \leq L} |t|^{-1} |R| dt$.

Let us show that $I_{2,1} \ll \mathcal{M} + \gamma_0$. For an infinitely many times differentiable function with bounded derivatives $h: \mathbf{R} \rightarrow \mathbf{C}$ write $c(h) = \|h'\| + \dots + \|h^{(\text{vi})}\|$. Let ξ be a standard normal random variable. By Lemma 3.1 of [7],

$$\left| \mathbf{E}h(S) - \mathbf{E}h(\xi) + \frac{N}{6} \mathbf{E}Y_1^3 \left(3\mathbf{E}h'(\xi) + 2\mathbf{E}h'''(\xi) \right) \right| \ll c(h)(\mathcal{M} + \gamma_0).$$

Choosing $h(x) = e\{itx\}$ we obtain from this bound that $|R| \ll (|t| + \dots + |t|^6)(\mathcal{M} + \gamma_0)$. Clearly, the last inequality implies $I_{2,1} \ll \mathcal{M} + \gamma_0$.

The proof of the bound $I_{2,2} \ll \mathcal{M} + \gamma_0$ is similar to that of Lemma 3.1 of [7], but somewhat more involved. Now the remainder $|R|$ of the expansion needs to be integrable with respect to the measure $dt/|t|$ over the region $\{|t| > c\}$. In order to construct such an estimate for the remainder we use the same argument as Bentkus and Götze [3]. That is, we approximate $\mathbf{E}e\{itS\}$ by the characteristic function of a statistic which is conditionally linear in the first m observations $X_1, \dots, X_m$ given the remaining part of the sample $X_{m+1}, \dots, X_n$. Here $m \approx Nt^{-2} \ln t$. The detailed calculations are given in [6].

REFERENCES

1. Babu J. G., Singh K. Inference on means using the bootstrap. — Ann. Statist., 1983, v. 11, p. 999–1003.
2. Bentkus V., Bloznelis M., Götze F. The Berry–Esseen bound for Student's statistic in the non-i.i.d. case. — J. Theoret. Probab., 1996, v. 9, № 3, p. 765–796.
3. Bentkus V., Götze F. The Berry–Esseen bound for Student's statistic. — Ann. Probab., 1996, v. 24, № 1, p. 491–503.
4. Bentkus V., Götze F., van Zwet W. R. An Edgeworth expansion for symmetric statistics. — Ann. Statist., 1997, v. 25, № 2, p. 851–896.
5. Bhattacharya R. N., Ghosh J. K. On the validity of the formal Edgeworth expansion. — Ann. Statist., 1978, v. 6, p. 435–451; Corrections *ibid.*, 1980, v. 8, p. 1399.
6. Bloznelis M., Putter H. Second order and bootstrap approximation to Student's t statistic. — Preprint 98-053 SFB 343. Bielefeld: Bielefeld University, 1998; <http://www.mathematik.uni-bielefeld.de/sfb343/Welcome.html>
7. Bloznelis M., Putter H. One term Edgeworth expansion for Student's t statistic. — Probability Theory and Mathematical Statistics: Proceedings of the Seventh Vilnius Conference, 1998. Ed. by B. Grigelionis et al. Vilnius, Utrecht: VSP/TEV, 1999, p. 81–98.

8. Чибисов Д. М. Асимптотическое разложение для распределения статистики, допускающей стохастическое разложение. I. — Теория вероятн. и ее примен., 1980, т. 25, в. 4, с. 745–756.
9. Chung K. L. The approximate distribution of Student's statistics. — Ann. Math. Statist., 1946, v. 17, p. 447–465.
10. Esseen C.-G. Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. — Acta Math., 1945, v. 77, p. 1–125.
11. Friedrich K. O. A Berry–Esseen bound for functions of independent random variables. — Ann. Statist., 1989, v. 17, № 1, p. 170–183.
12. Hall P. Edgeworth expansion for Student's t statistic under minimal moment conditions. — Ann. Statist., 1987, v. 15, p. 920–931.
13. Hall P. On the effect of random norming on the rate of convergence in the central limit theorem. — Ann. Probab., 1988, v. 16, № 3, p. 1265–1280.
14. Helmers R. On the Edgeworth expansion and the bootstrap approximation for a Studentized U -statistic. — Ann. Statist., 1991, v. 19, № 1, p. 470–484.
15. Helmers R. The Berry–Esseen bound for Studentized U -statistics. — Canad. J. Statist., 1985, v. 13, p. 79–82.
16. Helmers R., van Zwet W. R. The Berry–Esseen bound for U -statistics. — Statistical Decision Theory and Related Topics III: Proceedings of the Third Purdue Symposium, 1981, V. 1. Ed. by S. S. Gupta and J. O. Berger. New York: Academic Press, 1982, p. 497–512.
17. Prašková Z. Sampling from a finite set of random variables: the Berry–Esseen bound for the Studentized mean. — Proceedings of the Fourth Prague Symposium on Asymptotic Statistics. Prague: Charles Univ., 1989, p. 67–82.
18. Putter H., van Zwet W. R. Empirical Edgeworth expansions for symmetric statistics. — Ann. Statist., 1998, v. 26, № 4, p. 1540–1569.
19. Singh K. On the asymptotic accuracy of Efron's bootstrap. — Ann. Statist., 1981, v. 9, p. 1187–1195.
20. Slavova V. V. On the Berry–Esseen bound for Student's statistic. — Lecture Notes in Math., 1985, v. 1155, p. 335–390.

Поступила в редакцию
13.V.1999

© 2002 г.

GÖTZE F.*, TIKHOMIROV A. N.**

RATE OF CONVERGENCE TO THE SEMI-CIRCULAR LAW FOR THE GAUSSIAN UNITARY ENSEMBLE¹⁾

Показано, что расстояние Колмогорова между усредненным значением спектральной функции распределения матрицы Вигнера размера $n \times n$ с гауссовскими элементами и функцией распределения полукругового закона имеет порядок $O(n^{-2/3})$.

* Fakultät für Mathematik, Universität Bielefeld, 33501 Bielefeld 1, Germany; e-mail: goetze@mathematik.uni-bielefeld.de

** Сыктывкарский государственный университет, Октябрьский проспект., 55, 167001 Сыктывкар, Россия; e-mail: tikhomir@ssu.komi.com

¹⁾ Research of both authors supported by the DFG-Forschergruppe FOR 399/1-1 «Spektrale Analysis, Asymptotische Verteilungen und Stochastische Dynamiken». The second author partially supported by Russian Foundation for Basic Research (Grants № 99-01-00247, 99-01-00112, 00-15-96019) and by INTAS № 99-01317, DFG-RFBR № 99-01-04027.