A BERRY–ESSEEN BOUND FOR LEAST SQUARES ERROR VARIANCE ESTIMATORS OF REGRESSION PARAMETERS

M. Bloznelis   A. Račkauskas

ABSTRACT. We construct a precise Berry–Esseen bound for least squares error variance estimators of regression parameters. Our bound depends explicitly on the sequence of design variables and is of order $O(N^{-1/2})$ if this sequence is "regular" enough.

Key words: Berry–Esseen bound, least squares estimators, error variance, linear regression.

1. Introduction and results

Consider the linear model

$$Y_i = \alpha + \beta X_i + \varepsilon_i, \qquad i = 1, \dots, N, \tag{1.1}$$

where $Y_1, \dots, Y_N$ is the observed response to given variables (design points) $X = (X_1, \dots, X_N)$, and where α and β are unknown regression parameters to be estimated. Here $\varepsilon_1, \dots, \varepsilon_N$ are unobservable errors. We assume that $\varepsilon_1, \dots, \varepsilon_N$ are independent identically distributed mean zero random variables.

The ordinary least squares (OLS) estimators of the parameters α and β are defined by

$$\hat\alpha = \overline{Y} - \hat\beta\,\overline{X} \qquad \text{and} \qquad \hat\beta = \mathcal{V}^{-2}(x_1 y_1 + \dots + x_N y_N),$$

see Malinvaud (1970). Here $x_k = X_k - \overline{X}$, $y_k = Y_k - \overline{Y}$ and $\mathcal{V} = (x_1^2 + \dots + x_N^2)^{1/2}$.

We are interested in the rate of the normal approximation of $\hat\alpha - \alpha$ and $\hat\beta - \beta$ in the case where the variance $\sigma^2 = \mathbf{E}\varepsilon_1^2$ is unknown. In this case σ^2 is estimated by

$$s^2 = (\hat\varepsilon_1^2 + \dots + \hat\varepsilon_N^2)/(N-2), \qquad \hat\varepsilon_k = Y_k - \hat\alpha - \hat\beta\,X_k,$$

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX see Malinvaud (1970), and is used to normalize $\hat\alpha-\alpha$ and $\hat\beta-\beta$. By the central limit theorem, for large N, distributions of the statistics

$$\theta_1=\theta_1(X)=\frac{\hat\alpha-\alpha}{Q\,s}\,,\quad \theta_2=\theta_2(X)=\frac{\hat\beta-\beta}{\mathcal{V}\,s}\,,\quad \text{where}\quad Q=\Big(\frac{\overline{X}^2}{\mathcal{V}^2}+\frac{1}{N}\Big)^{1/2},$$

can be approximated by the standard normal distribution. We construct a bound for the rate of the normal approximation of these statistics.

The Berry–Esseen bound given in Theorem 1.1 below is precise and depends only on the ratio β_3/σ^3 and the fraction

$$\mathcal{E}=\mathcal{E}_X=\frac{\sum_{k=1}^N|x_k|^3}{\mathcal{V}^3}\,.$$

If $\mathcal{V}=0$ put $\mathcal{E}=1$ and $\theta_i=0$.

Theorem 1.1. *Assume that $\sigma^2>0$. Then there exists an absolute constant $c>0$ such that*

$$\sup_x\big|\mathbf{P}\{\theta_i\le x\}-\Phi(x)\big|\le c\,\mathcal{E}\,\frac{\beta_3}{\sigma^3}\,,\qquad i=1,2,$$

where $\Phi(x)$ denotes the standard normal distrubution function.

Here and below $c,c_0,c_1,\dots$ denote generic absolute constants. By $c(T_1,T_2,\dots)$ we denote a constant which depends only on the quantities indicated in the brackets.

If the sequence X is regular enough, e.g., the design points are regularly spaced, $X_k=a+k\,t$, $k=1,\dots,N$, for some a, t, then Theorem 1.1 yields the Berry–Esseen bound

$$\sup_x\big|\mathbf{P}\{\theta_i\le x\}-\Phi(x)\big|\le c\,\frac{\beta_3}{\sqrt{N}\,\sigma^3}\,,\qquad \text{for}\qquad i=1,2.$$

The rate $O(N^{-1/2})$ is achieved also when the design points are taken at random and independently of the errors. Consider the linear model (1.1) in the case where the design points $\mathcal{X}=(\mathcal{X}_1,\dots,\mathcal{X}_N)$ are i.i.d. random variables independent of $\varepsilon_1,\dots,\varepsilon_N$. Write $\gamma_3=\mathbf{E}\,|\mathcal{X}_1-\mathbf{E}\,\mathcal{X}_1|^3$ and $\tau^2=\operatorname{Var}\mathcal{X}_1$. Theorem 1.1 implies the following bound.

Corollary 1.1. *Assume that $\sigma^2>0$ and $\tau^2>0$. Then*

$$\sup_x\big|\mathbf{P}\{\theta_i(\mathcal{X})\le x\}-\Phi(x)\big|\le c\,\frac{\beta_3\,\gamma_3}{\sqrt{N}\sigma^3\,\tau^3}\,,\qquad \textit{for}\qquad i=1,2.$$

It is easy to recognize the structural similarity between the statistics θ_1 and θ_2 and the Student statistics, based on non-identically distributed observations. The Berry-Esseen bound for the Student test was proved by Bentkus and Götze (1996) and extended to the non-i.i.d. case in Bentkus, Bloznelis and Götze (1996) ([BBG] for short). In order to obtain the optimal results for error variance estimators of regression parameters we extend the methods developed in these papers and then apply them to the statistics θ_1 and θ_2.

It seems that the traditional technique, see, e.g., Lahiri (1992), does not allow to obtain the optimal results: it involves moments of order which is higher than the optimal one.

The rest of the paper is divided in two sections. In Section 2 we present a general Berry–Esseen bound for the statistics like θ_1 and θ_2 which is obtained without the assumption that the second moment of the errors ε_i is finite. Theorem 1.1 is a simple consequence of this result. Proofs are given in Section 3.

2. A general result

Given two sequences of real numbers $w = \{w_1, \dots, w_N\}$ and $z = \{z_1, \dots, z_N\}$ such that

$$\sum_{i=1}^N w_i^2 = 1 \qquad \text{and} \qquad \sum_{i=1}^N z_i^2 = 1 \tag{2.1}$$

write

$$T = T(w,z) = S(w)/s(z), \qquad S(w) = \sum_{i=1}^N w_i \varepsilon_i, \qquad s^2(z) = \hat\sigma^2 - N^{-1} S^2(z),$$

where as usual

$$\overline{\varepsilon} = \big(\varepsilon_1 + \cdots + \varepsilon_N\big)/N \qquad \text{and} \qquad \hat\sigma^2 = N^{-1} \sum_{i=1}^N \big(\varepsilon_i - \overline{\varepsilon}\big)^2.$$

Write $\mathcal{E}(w) = \sum_{i=1}^N |w_i|^3$ and note that by (2.1),

$$\mathcal{E}(w) \ge N^{-1/2} \qquad \text{and} \qquad \sum_{i=1}^N |w_i| \le N^{1/2}. \tag{2.2}$$

Clearly the same inequalities hold for $\mathcal{E}(z)$ and $\sum_{i=1}^N |z_i|$ as well.

Define the number $a^2 = a_N^2$ by the truncated second moment equation,

$$a^2 = \sup\big\{b : \ \mathbf{E}\, \varepsilon_1^2 \mathbf{1}\{\varepsilon_1^2 \le b\, N\} \ge b\big\}, \qquad a \ge 0.$$

It is known (and easy to show) that for any random variable ε_1 and any N such a number a exists, and a is the largest solution of the equation $a^2 = \mathbf{E}\, \varepsilon_1^2 \mathbf{1}\{\varepsilon_1^2 \le a^2 N\}$. Moreover, if $\sigma < \infty$ then $a \le \sigma$. If $\sigma = \infty$ then $a_N \to \infty$ as $N \to \infty$.

In order to formulate the general result we introduce truncated random variables

$$\tau_i = a^{-1}N^{-1/2}\varepsilon_i \mathbf{1}\{\varepsilon_i^2 \le a^2 N\}, \qquad \text{for} \quad 1 \le i \le N.$$

Note that $|\tau_i| \le 1$ and $\mathbf{E}\,\tau_i^2 = 1/N$.

The result formulated below provides a bound for

$$\delta_N = \sup_x \bigl|\mathbf{P}\{T \le x\} - \Phi(x)\bigr|$$

in the general case, where the existence of the second moment of r.v. ε_1 is not assumed.

Theorem 2.1. *Assume that $a > 0$. Then*

$$\delta_N \le c\,N\,\mathbf{P}\{\varepsilon_1^2 > a^2 N\} + c\,N\,\bigl|\mathbf{E}\,\tau_1\bigr| + c\,N^{3/2}\bigl(\mathcal{E}(w) + \mathcal{E}(z)\bigr)\,\mathbf{E}\,|\tau_1|^3. \tag{2.3}$$

3. Proofs

Proof of Theorem 1.1. A simple calculation shows that $\hat{\theta}_2 = (1-2/N)^{1/2}T(z,z)$ and $\hat{\theta}_1 = (1-2/N)^{1/2}T(w,z)$, where $z = \{z_1,\dots,z_N\}$, $z_i = x_i/\mathcal{V}$ and where

$$w = \{w_1,\dots,w_N\}, \quad w_i = v_i/W, \quad v_i = N^{-1} - x_i\overline{X}/\mathcal{V}^2, \quad W^2 = \overline{X}^2/\mathcal{V}^2 + N^{-1}.$$

By Theorem 2.1 and inequalities (2.2),

$$\sup_x \bigl|\mathbf{P}\{\theta_i \le x\} - \Phi(x)\bigr| \le c\,\frac{\beta_3}{\sigma^3}\,\bigl(\mathcal{E}(w) + \mathcal{E}(z)\bigr), \qquad \text{for} \quad i = 1,2.$$

Furthermore, by Hölder's inequality, $\overline{X}^2 \le \mathcal{V}^2/N$. Therefore,

$$\sum_{i=1}^N |v_i|^3 \le c\,N^{-2} + c\,(|\overline{X}|/\mathcal{V}^2)^3 \sum_{i=1}^N |x_i|^3 \le c\,N^{-3/2}\bigl(N^{-1/2} + \mathcal{E}(z)\bigr)$$

and we obtain $\mathcal{E}(w) \le c\mathcal{E}(z)$. But $\mathcal{E}(z) = \mathcal{E}$ and this completes the proof.

Proof of Corollary 1.2. The result follows from Theorem 1.1 and the inequality

$$\mathbf{E}\,\mathcal{E}_{\mathcal{X}} \le c\,\frac{\gamma_3}{\sqrt{N}\,\tau^3}\,.$$

Let us prove this inequality. We asume without loss of generality that $\mathbf{E}\,\mathcal{X}_1 = 0$. Write $\chi_i = \mathcal{X}_i - \overline{\mathcal{X}}$, for $i = 1,\dots,N$, where $\overline{\mathcal{X}} = N^{-1}(\mathcal{X}_1 + \cdots + \mathcal{X}_N)$, and $v_s = \mathbf{E}\,|\chi_1|^s$ and $L = v_3\,v_2^{-3/2}$. Denote the event

$$\mathcal{A} = \bigl\{\chi \le \mathbf{E}\,\chi/2\bigr\}, \qquad \text{where} \qquad \chi = \chi_1^2 + \cdots + \chi_N^2.$$

Note that $\mathcal{E}_{\mathcal{X}} \le 1$ and thus, $\mathbf{E}\,\mathcal{E}_{\mathcal{X}} \le P\{\mathcal{A}\} + 2^{3/2}N^{-1/2}\,L$. By Chebyshev's inequality and Rosenthal's inequality, see, e.g. Petrov (1995),

$$\mathbf{P}\{\mathcal{A}\} \le c\mathbf{E}\,|\chi - \mathbf{E}\,\chi|^{3/2}/(N\,v_2)^{3/2} \le c\,N^{-1/2}\,L.$$

We obtain $\mathbf{E}\,\mathcal{E}_{\mathcal{X}} \le c\,N^{-1/2}L$. Finally, using the inequalities

$$\mathbf{E}\,\chi_1^2 = (1 - N^{-1})\,\tau^2 \ge \tau^2/2 \qquad \text{and} \qquad \mathbf{E}\,|\chi_1|^3 \le c\,\gamma_3.$$

we obtain $L \le c\,\gamma_3/\tau^3$ thus completing the proof.

Proof of Theorem 2.1. The proof goes along the lines of the proof of Theorem 1.2 in [BBG]. Therefore we give only a general scheme of the proof. These steps of the proof which are new are presented with detail calculations.

Write

$$\varkappa_i = w_i(\tau_i - \mathbf{E}\,\tau_i), \qquad \eta_i = \tau_i^2 - N^{-1}, \qquad 1 \le i \le N,$$
$$\varkappa = \sum_{i=1}^N \varkappa_i, \qquad \tau = \sum_{i=1}^N \tau_i \qquad \text{and} \qquad \eta = \sum_{i=1}^N \eta_i.$$

Let $\alpha = (\alpha_1, \dots, \alpha_N)$ denote a sequence of i.i.d. Bernoulli random variables such that $\mathbf{P}\{\alpha_1 = 1\} = 1 - \mathbf{P}\{\alpha_1 = 0\} = p$, with some $p \le 1$. By ξ we shall denote a standard normal random variable. By (2.1), we can write $\xi = \xi_1 + \dots + \xi_N$, where $\xi_1, \dots, \xi_N$ are independent centered normal random variables with $\mathbf{E}\,\xi_i^2 = w_i^2$, $i = 1, 2, \dots, N$. We shall assume that the sequences α, $(\tau_1, \dots, \tau_N)$ and $(\xi_1, \dots, \xi_N)$ are independent. Furthermore, we may and shall assume that

$$N\,\bigl|\mathbf{E}\,\tau_1\bigr| < c_0 \qquad \text{and} \qquad \bigl(\mathcal{E}(w) + \mathcal{E}(z)\bigr)\,N^{3/2}\mathbf{E}\,|\tau_1|^3 < c_0, \tag{3.1}$$

for a sufficiently small $c_0 > 0$. If at least one of these inequalities fails then (2.3) follows from the obvious estimate $\delta_N \le 1$. Note that the second inequality in (3.1) implies $N\,\mathbf{E}\,|\tau_1|^3 < c_0$.

Let $g : R \to R$ denote a function which is infinitely many times differentiable with bounded derivatives and such that

$$1/8 \le g(x) \le 2, \quad \text{for all} \quad x \in R, \quad \text{and} \quad g(x) = |x|^{-1/2}, \quad \text{for} \quad 1/4 \le |x| \le 7/4.$$

For a statistic W write $\delta(W) = \sup_x\bigl|\mathbf{P}\{W \le x\} - \Phi(x)\bigr|$.

The proof is divided in two steps. In the first step we replace the statistic T by the statistics $V = \sqrt{N}\varkappa g(1+\eta)$, which is a smooth function of the random variables $\varkappa_i$ and η_i. In the second step we construct estimates for the difference between the characteristic functions $f(t) = \mathbf{E}\,\exp\{i\,t\,V\}$ and $\phi(t) = \mathbf{E}\,\exp\{i\,t\,\xi\} = \exp\{-t^2/2\}$

and then give a bound for $\delta(V)$ using Esseen's smoothing lemma, see, e.g., Feller (1971).

Step 1. We shall show that

$$\delta(T) \le \delta(V) + c\,\mathcal{R}_1, \quad \text{where} \quad \mathcal{R}_1 = N\,P\{\varepsilon_1^2 > a^2 N\} + N\,|\mathbf{E}\,\tau_1| + N\,\mathbf{E}\,|\tau_1|^3. \tag{3.2}$$

Write $V' = \sqrt{N}\,S_1\,g(1+A)$ and $V'' = \sqrt{N}\varkappa g(1+A)$, where

$$S_1 = \sum_{i=1}^{N} w_i \tau_i, \qquad S_2 = \sum_{i=1}^{N} z_i \tau_i, \qquad \text{and} \qquad A = \eta + \tau^2/N + S_2^2.$$

Let us show that $\delta(T) \le \delta(V') + c\,\mathcal{R}_1$. We have

$$\mathbf{P}\{T \ne V'\} \le N\,\mathbf{P}\{\varepsilon_1^2 \ge a^2 N\} + \mathbf{P}\{|A| \ge 3/4\}.$$

By Chebyshev's inequality, Lemma 3.1 (see below) and the inequality $N^{-1/2} \le N\,\mathbf{E}\,|\tau_1|^3$,

$$\mathbf{P}\{|\eta| \ge 1/4\} \le c\,\mathbf{E}\,|\eta|^{3/2} \le c\,\mathcal{R}_1, \qquad \mathbf{P}\{\tau^2/N \ge 1/4\} \le c\,N^{-1} \le c\,\mathcal{R}_1,$$

and $\mathbf{P}\{S_2^2 \ge 1/4\} \le c\,\mathbf{E}\,S_2^2$. Furthermore, by (2.2) and (3.1),

$$\mathbf{E}\,S_2^2 \le \mathbf{E}\,\tau_1^2 + |\mathbf{E}\,\tau_1|^2 \sum_{i \ne j} |z_i|\,|z_j| \le N^{-1} + N\,|\mathbf{E}\,\tau_1|^2 \le \mathcal{R}_1.$$

We obtain $\mathbf{P}\{|A| > 3/4\} \le c\,\mathcal{R}_1$ and therefore $\big|\delta(T) - \delta(V')\big| \le c\,\mathcal{R}_1$.

Furthermore, $|V' - V''| \le 2\,R_1(w)$, where $R_1(w) = \sqrt{N}\,|\mathbf{E}\,\tau_1| \sum_{i=1}^N |w_i|$. It follows from (2.2) that $R_1(w) \le R_1 := N\,|\mathbf{E}\,\tau_1|$ and thus $\delta(V') \le \delta(V'') + c\,\mathcal{R}_1$.

It remains to show that $\delta(V'') \le \delta(V) + c\,\mathcal{R}_1$. Expanding g in powers of $Q = N^{-1}\tau^2 + S_2^2$ we obtain $V'' = V + R_2$, where $|R_2| \le c\sqrt{N}|\varkappa Q|$. By Chebyshev's inequality,

$$\mathbf{P}\{|R_2| \ge N^{-1/2}\} \le c\,N\,\mathbf{E}\,|\varkappa Q| \le c\,\big(\mathbf{E}\,|\varkappa|^3\big)^{1/3}\big(\mathbf{E}\,|\tau|^3\big)^{2/3} + c\,N\,\mathbf{E}\,|\varkappa| S_2^2. \tag{3.3}$$

By Lemma 3.1, the first summand is bounded by $c\,N^{-1/2}$. To estimate the second one let us write $u = \sum_{i=1}^N z_i\,(\tau_i - \mathbf{E}\,\tau_i)$ and $R_1(z) = \sqrt{N}\,|\mathbf{E}\,\tau_1| \sum_{i=1}^N |z_i|$. Clearly, $\mathbf{E}\,|N^{1/2}u|^3 \le c$, cf. Lemma 3.1, and $R_1(z) \le R_1$. Since $N S_2^2 \le 2\,N\,u^2 + 2\,R_1^2(z)$ and $|\varkappa| R_1^2(z) \le |\varkappa|^3 + |R_1|^3$ and $u^2|\varkappa| \le |u|^3 + |\varkappa|^3$, the last summand in (3.3) is bounded by $cN^{-1/2} + c\,R_1^3$. We obtain $\mathbf{P}\{|R_2| \ge N^{-1/2}\} \le c\,\mathcal{R}_1$. Therefore, $\delta(V'') \le \delta(V) + c\,\mathcal{R}_1$. We arrive to (3.2).

Step 2. Here we show that

$$\delta(V) \le c\mathcal{R}_2, \qquad \mathcal{R}_2 = N|\mathbf{E}\,\tau_1| + N^{3/2}\mathcal{E}(w)\,\mathbf{E}\,|\tau_1|^3. \tag{3.4}$$

It follows from Esseen's smoothing lemma that

$$\delta(V) \le \int_{0\le|t|\le T_1} |t|^{-1}|f(t)-\phi(t)|dt + \frac{c}{T_1}\,, \quad T_1 = c_2/(N^{3/2}\mathcal{E}(w)\,\mathbf{E}\,|\tau_1|^3).$$

This inequality implies (3.4) if we show that

$$I_1 := \int_{0\le|t|\le c_1} |t|^{-1}|f(t)-\phi(t)|dt \le c\mathcal{R}_2 \tag{3.5}$$

and

$$I_2 := \int_{c_1\le|t|\le T_1} |t|^{-1}|f(t)-\phi(t)|dt \le c\mathcal{R}_2, \tag{3.6}$$

where the constant c_1 is sufficiently large. Let us prove (3.5). Write $H(x) = \exp\{itx\}$ and denote $c_H = \|H'\| + \|H''\| + \|H'''\|$, where $\|H\| = \sup_x |H(x)|$. Then (3.5) follows from the inequality

$$\big|\mathbf{E}\,H(V) - \mathbf{E}\,H(\xi)\big| \le c\,c_H\,\mathcal{R}_2. \tag{3.7}$$

This inequality is a consequence of

$$\big|\mathbf{E}\,H(\sqrt{N}\varkappa) - \mathbf{E}\,H(\xi)\big| \le c\,c_H\,\mathcal{R}_2, \quad \big|\mathbf{E}\,H(V) - \mathbf{E}\,H(\sqrt{N}\varkappa)\big| \le c\,c_H\,\mathcal{R}_2. \tag{3.8}$$

The proof of the first inequality in (3.8) is easy, see Bentkus, Götze, Paulauskas and Račkauskas (1990). The proof of the second inequality is very close to the proof of the inequality (2.3) in Bentkus and Götze (1996). An inspection of their proof shows that in order to verify the second inequality of (3.8) it suffices to show that

$$\mathbf{E}\,|D| \le c\,\mathcal{R}_2, \quad \mathbf{E}\,U^2 \le c\,\mathcal{R}_2, \quad \mathbf{E}\,\eta^2 \le c\,\mathcal{R}_2, \quad \big|\mathbf{E}\,U\,H'(\sqrt{N}\,\varkappa)\big| \le c\,\mathcal{R}_2, \tag{3.9}$$

where $D = \sqrt{N}\sum_{i=1}^N \varkappa_i\,\eta_i$ and $U = \sqrt{N}\sum_{1\le i,j\le N,\, i\ne j} \varkappa_i\eta_j$. The first two inequalities of (3.9) follows from Lemma 3.1. The inequality $\mathbf{E}\,\eta^2 \le c\,\mathcal{R}_2$ is trivial, since $|\eta_i| \le 2$, for $1 \le i \le N$. To prove the last inequality in (3.9) let us write

$$U\,H'(\sqrt{N}\,\varkappa) = \sum_{1\le i,j\le N,\, i\ne j} A_{i,j}, \qquad A_{i,j} = \sqrt{N}\,\varkappa_i\eta_j H'(\sqrt{N}\,\varkappa)$$

and $\varkappa = \varkappa^{i,j} + \varkappa_i + \varkappa_j$, where we denote $\varkappa^{i,j} = \varkappa - \varkappa_i - \varkappa_j$. Expanding $H'\big(\sqrt{N}\,(\varkappa^{i,j}+\varkappa_i+\varkappa_j)\big)$ in powers of $\sqrt{N}\,\varkappa_i$ and subsequently in powers of $\sqrt{N}\,\varkappa_j$ we get

$$\big|\mathbf{E}A_{i,j}\big| \le c\,\mathbf{E}(\sqrt{N}\varkappa_i)^2\big|\sqrt{N}\,\varkappa_j\eta_j\big| \le c\,w_i^2|w_j|N^{1/2}\mathbf{E}\,|\tau_1|^3.$$

An application of (2.2) completes the proof of (3.9). We arrive to (3.7)

The proof of the (3.6) is more complicated. An additional effort has to be made to ensure the integrability (with respect to the measure $dt/|t|$ in the region $c_1 \le |t| \le T_1$) of the remainder of expansions. In order to prove (3.6) one uses the approach developed in Callaert and Janssen (1978), Helmers and van Zwet (1982), van Zwet (1984) where similar integral is estimated in the case where $f(t)$ is the characteristic function of U or more general nonlinear symmetric statistic based on i.i.d. observations. One extension of this approach to the situation where the observations are *non-identically distributed* is given in [BBG].

To prove (3.6) we expand g in powers of observations and then expand the exponent in much the same way as in proof of Theorem 1.2 of [BBG], only now we use inequalities of Lemma 3.1 to bound the remainders of these expansions.

Finally, combining (3.5) and (3.6) we obtain (3.4) thus completing the proof of the theorem.

Lemma 3.1. *Assume that* (3.1) *holds. Then*

$$\mathbf{E}\,\big|\tau_1+\cdots+\tau_k\big|^s \le c(s), \quad \textit{for} \quad s>0, \quad \textit{and} \quad k=1,2,\dots,N,$$
$$\mathbf{E}\,\big|\sqrt{N}\sum_{i=1}^N \alpha_i\varkappa_i\eta_i\big| \le c\,p\,N\,\mathbf{E}\,|\tau_1|^3, \qquad \mathbf{E}\,\big|\sum_{i=1}^N \alpha_i\eta_i\big|^{3/2} \le c\,p\,N\,\mathbf{E}\,|\tau_1|^3,$$
$$\mathbf{E}\,\big|\sqrt{N}\,\varkappa\big|^3 \le c, \quad \mathbf{E}\,U^2 \le c\,p^2\,N\,\mathbf{E}\,|\tau_1|^3, \quad \mathbf{E}\,|U|^{3/2} \le c\,p^{7/4}\,N\,\mathbf{E}\,|\tau_1|^3,$$

where $U = \sqrt{N}\sum_{1\le i,j\le N,\, i\ne j}\alpha_i\varkappa_i\alpha_j\eta_j$.

Proof of Lemma 3.1. The proof is easy and routine, cf. proof of Lemma 2.1 in [BBG].

References

Bentkus, V., Götze, F., *The Berry–Esséen bound for Student's statistic*, Ann. Probab. **24** (1996), 491–503.

Bentkus, V. Bloznelis, M. and Götze, F., *A Berry–Esséen bound for Student's statistic in the non-i.i.d. case*, J. Theoret. Probab. **9** (1996), 765–796.

Bentkus V., Götze F., Paulauskas V. and Račkauskas A., *The accuracy of Gaussian approximation in Banach spaces*, SFB 343 Preprint 90 -100, Universität Bielefeld (In Russian: Itogi Nauki i Tehniki, ser. Sovr. Probl. Matem., Moskva, VINITI **81** (1991), 39–139; to appear in English in Encyclopedia of Mathematical Sciences, Springer), 1990.

Callaert, H. and Janssen, P., *The Berry–Esseen theorem for U-statistics*, Ann. Statist. **6** (1978), 417–421.

Feller, W., *Probability Theory, Vol 2*, Wiley, New York, 1971.

Helmers, R., and van Zwet, W. R., *The Berry–Esseen bound for U-statistics.* **1** (1982), Statistical Decision Theory and Related Topics,III (S.S. Gupta and J.O. Berger, eds), Academic Press, New York,, 497–512.

Lahiri, S. N., *Bootstrapping M-estimators of a multiple linear regression parameter.*, Ann. Statist. **20** (1992), 1548–1570.

Malinvaud, E., *Statistical Methods of Econometrics*, North–Holland, Amsterdam, 1970.

van Zwet, W.R., *A Berry-Esseen bound for symmetric statistics*, Z. Wahrsch. verw. Gebiete **66** (1984), 425–440.