Centrum voor Wiskunde en Informatica

REPORT *RAPPORT*

PNA

Probability, Networks and Algorithms

Probability, Networks and Algorithms

PNA

The empirical Edgeworth expansion for a studentized trimmed mean

N. Gribkova, R. Helmers

REPORT PNA-R0214 MAY 31, 2002

CWI is the National Research Institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organization for Scientific Research (NWO).
CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

Information Systems (INS)

P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

ISSN 1386-3711

The Empirical Edgeworth Expansion for a Studentized Trimmed Mean

Nadezhda Gribkova
St.Petersburg State Transport University
190031, Moskovskii av. 9, St.Petersburg, Russia
E-mail: nvg@spb.cityline.ru

Roelof Helmers
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
E-mail: R.Helmers@cwi.nl

ABSTRACT

We establish the validity of the empirical Edgeworth expansion (EE) for a studentized trimmed mean, under the sole condition that the underlying distribution function of the observations satisfies a local smoothness condition near the two quantiles where the trimming occurs. A simple explicit formula for the $n^{-1/2}$ term (correcting for skewness and bias; n being the sample size) of the EE will be given. In particular our result supplements previous work by Hall and Padmanabhan (1992) and Putter and van Zwet (1998). The proof is based on a U-statistic type approximation and also uses a version of Bahadur's (1966) representation for sample quantiles.

2000 Mathematics Subject Classification: Primary 62E20, 62G30; Secondary 60F05.
Keywords and Phrases: Empirical Edgeworth expansions, studentized trimmed means.
Note: Work carried out under project PNA 3.2 Statistics.

1. Introduction and main results

In this paper we establish the validity of the empirical Edgeworth expansion (EE) for a studentized trimmed mean, under the sole condition that underlying distribution function (*df*) of the observations satisfies a local smoothness condition near the two quantiles where the trimming occurs. In particular our result supplements previous work by Hall and Padmanabhan (1992) and Putter and van Zwet (1998). The existence of an Edgeworth expansion (EE) for a studentized trimmed mean was also obtained by Hall and Padmanabhan (1992), but these authors wrote that the "first term in an Edgeworth expansion is very complex and so it will not be written down explicitly". In contrast, in the present paper we show that our method of proof gives a simple explicit formula for the $N^{-1/2}$- term (correcting for skewness and bias; N being the sample size) of the Edgeworth expansion. The proof of our result is based on a $U-$statistic type approximation (cf. also Bickel, Götze, van Zwet (1986) and Helmers (1991)) and also uses a version of Bahadur's (1966) representation for sample quantiles. We will also show (cf. Lemma A.2, Appendix) that our result cannot be obtained as a consequence of a general result (Theorem 1.2 of [10]) for a studentized symmetric statistics of Putter and van Zwet (1998).

Let $X_1, \ldots, X_N$ be independent and identically distributed (i.i.d.) real-valued random variables (r.v.) with common *df* F, and let $X_{1:N} \leq \cdots \leq X_{N:N}$ denote the corresponding order statistics. Consider the trimmed mean given by

$$T_N = \frac{1}{([\beta N] - [\alpha N])} \sum_{i=[\alpha N]+1}^{[\beta N]} X_{i:N} \ , \tag{1.1}$$

where $0 < \alpha < \beta < 1$ are any fixed numbers and $[\cdot]$ represents the greatest integer function. Let $F^{-1}(u) = \inf\{x : F(x) \geq u\}$, $0 < u \leq 1$, denote the left-continuous inverse function of *df* F and put $\frac{d}{du}F^{-1}(u) = 1/f(F^{-1}(u))$ to be its derivative, when the density $f = F'$ exists and $f(F^{-1}(u)) > 0$. Let

$$\xi_\nu = F^{-1}(\nu),$$

$0 < \nu < 1$, be the ν-th quantile of F. Define a function

$$Q(u) = \begin{cases} \xi_\alpha\ , & u \leq \alpha, \\ F^{-1}(u)\ , & \alpha < u \leq \beta, \\ \xi_\beta\ , & \beta < u. \end{cases}$$

Let W_i, $i = 1, \ldots, N$, denote X_i Winsorized outside of $(\xi_\alpha, \xi_\beta]$, that is

$$W_i = \begin{cases} \xi_\alpha\ , & X_i \leq \xi_\alpha, \\ X_i\ , & \xi_\alpha < X_i \leq \xi_\beta, \\ \xi_\beta\ , & \xi_\beta < X_i. \end{cases} \tag{1.2}$$

Then $W_i \overset{\mathrm{d}}{=} Q(U_i)$, $i = 1, \ldots, N$, where U_i are independent r.v.'s with the uniform on $(0, 1)$ distribution. Define

$$\mu_W = \int_0^1 Q(u)\, du, \quad \sigma_W^2 = \int_0^1 (Q(u) - \mu_W)^2\, du, \quad \gamma_{3,W} = \int_0^1 (Q(u) - \mu_W)^3\, du. \tag{1.3}$$

Put

$$\delta_{2,W} = -\alpha^2 \frac{1}{f(\xi_\alpha)}[\mu_W - \xi_\alpha]^2 + (1-\beta)^2 \frac{1}{f(\xi_\beta)}[\mu_W - \xi_\beta]^2. \tag{1.4}$$

Suppose that $\xi_\alpha \neq \xi_\beta$ (that is ξ_α is not an atom with mass $(\beta - \alpha)$ for the distribution F), then W_i are not degenerated ($\sigma_W > 0$). Define two real numbers λ_1 and λ_2 as

$$\lambda_1 = \gamma_{3,W}/\sigma_W^3, \qquad \lambda_2 = \delta_{2,W}/\sigma_W^3. \tag{1.5}$$

We need no moment assumptions about the distribution F and to normalize T_N we use

$$\mu(\alpha, \beta) = \frac{1}{\beta - \alpha} \int_\alpha^\beta F^{-1}(u)\, du \tag{1.6}$$

as a location parameter and $(\beta - \alpha)^{-1}\sigma_W$ (the root of the asymptotic variance) as a scale parameter. Note that T_N often serves as a statistical estimator for the parameter $\mu(\alpha, \beta)$.

Now we show why moments are not needed. Take some fixed $\Delta > 0$ and define auxiliary i.i.d. Winzorized r.v.'s $X_i' = \max(\xi_\alpha - \Delta, \min(X_i, \xi_\beta + \Delta))$. Let $X_{i:N}'$, $i = 1, \ldots, N$, denote the correspondent order statistics. Introduce auxiliary trimmed mean $T_N' = \frac{1}{([\beta N] - [\alpha N])} \sum_{i=[\alpha N]+1}^{[\beta N]} X_{i:N}'$. Fix an arbitrary $x \in R$ and note that

$$\{T_N \leq x\} \subset \{T_N' \leq x\} \cup \{X_{[\alpha N]+1:N} < \xi_\alpha - \Delta\} \cup \{X_{[\beta N]:N} > \xi_\beta + \Delta\},$$

$$\{T_N' \leq x\} \subset \{T_N \leq x\} \cup \{X_{[\alpha N]+1:N} < \xi_\alpha - \Delta\} \cup \{X_{[\beta N]:N} > \xi_\beta + \Delta\}.$$

If the conditions of our Theorem 1.1 are satisfied then, by Bernstein's inequality we get $P(X_{[\alpha N]+1:N} < \xi_\alpha - \Delta) + P(X_{[\beta N]:N} > \xi_\beta + \Delta) = O(exp(-cN))$, as $N \to \infty$, where $c > 0$ is some independent on N constant. Therefore

$$\sup_{x \in R} |P(T_N \leq x) - P(T_N' \leq x)| = O(e^{-cN}) \tag{1.7}$$

and when proving our results we can replace with impunity T_N by T_N', which has finite moments of the arbitrary order.

In absence of any moment assumptions, our formula for the first $N^{-1/2}$ term of EE contains a bias term. Define a quantity

$$\beta_N = \frac{1}{N}\Big\{-(\alpha N - [\alpha N])\Big(\mu(\alpha,\beta)-\xi_\alpha\Big) - \frac{1}{2}\alpha(1-\alpha)\frac{1}{f(\xi_\alpha)}$$
$$+(\beta N - [\beta N])\Big(\mu(\alpha,\beta)-\xi_\beta\Big) + \frac{1}{2}\beta(1-\beta)\frac{1}{f(\xi_\beta)}\Big\}. \tag{1.8}$$

Note that when αN and βN are integer valued, the bias term has peculiary simple form: $\beta_N = \frac{1}{2N}\Big\{-\frac{\alpha(1-\alpha)}{f(\xi_\alpha)} + \frac{\beta(1-\beta)}{f(\xi_\beta)}\Big\}$. Moreover, in case $\alpha = 1-\beta$ and $f(\xi_\alpha) = f(\xi_\beta)$ (when the distribution F is symmetric, for example), the bias term vanishes.

We show (cf. Lemma A.1, Appendix) that if the conditions of our Theorem 1.1 are satisfied, then for an arbitrary $\Delta > 0$

$$b_N = (\beta-\alpha)(ET'_N - \mu(\alpha,\beta)) = \beta_N + O(N^{-3/2}) \tag{1.9}$$

as $N \to \infty$.(cf. (1.7)) Note also that the bias term (1.8) does not depend on Δ.

Define

$$F_{T_N}(x) = P\left(\frac{N^{1/2}(T_N - \mu(\alpha,\beta))}{(\beta-\alpha)^{-1}\sigma_W} \le x\right) \tag{1.10}$$

to be a *df* of a normalized trimmed mean. The Edgeworth expansion for the distribution function $F_{T_N}(x)$ is given by

$$G_N(x) = \Phi(x) - \frac{\phi(x)}{6\sqrt{N}}\left((\lambda_1 + 3\lambda_2)(x^2-1) + 6N\frac{\beta_N}{\sigma_W}\right), \tag{1.11}$$

where Φ is the standard normal distribution function, $\phi = \Phi'$. We use here the notation of Putter and van Zwet (1998). The quantity $(\lambda_1 + 3\lambda_2)N^{-1/2}$ serves as an approximation to the third cumulant of $\frac{N^{1/2}(T_N - \mu(\alpha,\beta))}{(\beta-\alpha)^{-1}\sigma_W}$, moreover $\lambda_1 N^{-1/2}$ is the approximation to the third cumulant of the L_2-projection of the normalized trimmed mean, which close to $N^{-1/2}\sigma_W^{-1}\sum_1^N W_i$ - sum of N i.i.d. Winsorized r.v. (cf. Sect.3, below), and $3\lambda_2 N^{-1/2}$ is the correcting term (note that in the case of trimmed mean $3\lambda_2 = 3\delta_{2,W}\sigma_W^{-3} = \alpha^2 h'(\alpha) - (1-\beta)^2 h'(\beta)$, where $h(u) = \big((\mu_W - F^{-1}(u))/\sigma_W\big)^3$).

Here is our first result: an Edgeworth expansion for a normalized trimmed mean.

THEOREM 1.1. *Suppose that $f = F'$ exists in neighborhoods of the points ξ_α and ξ_β, satisfies a Lipschitz condition there and $f(\xi_\nu) > 0$, $\nu = \alpha, \beta$. Then*

$$\sup_{x\in R} |F_{T_N}(x) - G_N(x)| = o(N^{-1/2}), \tag{1.12}$$

as $N \to \infty$.

Theorem 1.1 can be viewed as a version of the Edgeworth expansion for trimmed mean obtained by Bjerve (1974) in his unpublished Berkeley Ph.D. thesis (cf. also Helmers (1979)). Our method of proof is different from Bjerve's, as he used a conditioning argument to reduce a trimmed mean to a sum of i.i.d. r.v.'s, conditionally given the values of $X_{[\alpha N]+1,N}$ and $X_{[\beta N],N}$, while in contrast we essentially show that T_N can be approximated by a $U-$statistic U_N; the remainder $T_N - U_N$ can be shown to be of negligible order for our purposes by an application of a version of Bahadur (1966) representation for sample quantiles.

Next we state our result on the validity of one-term Edgeworth expansion for the Studentized trimmed mean. Define plug-in estimators for μ_W and σ_W^2 by

$$\hat{\mu}_W = \frac{k}{N} X_{k:N} + \frac{1}{N} \sum_{i=k+1}^{m-1} X_{i:N} + \frac{N-m+1}{N} X_{m:N}, \tag{1.13}$$

and

$$S_N^2 = \left(\frac{k}{N} X_{k:N}^2 + \frac{1}{N} \sum_{i=k+1}^{m-1} X_{i:N}^2 + \frac{N-m+1}{N} X_{m:N}^2 \right) - \hat{\mu}_W^2 \tag{1.14}$$

with $k = [\alpha N] + 1$ and $m = [\beta N]$. Let

$$F_{N,S}(x) = P\left(\frac{N^{1/2}(T_N - \mu(\alpha, \beta))}{(\beta - \alpha)^{-1} S_N} \le x \right) \tag{1.15}$$

denote a *df* of a studentized trimmed mean. Define

$$H_N(x) = \Phi(x) + \frac{\phi(x)}{6\sqrt{N}} \left((2x^2+1)\lambda_1 + 3(x^2+1)\lambda_2 - 6N \frac{\beta_N}{\sigma_W} \right). \tag{1.16}$$

Our main result is:

THEOREM 1.2. *Suppose that the conditions of Theorem 1.1 are satisfied. Then*

$$\sup_{x \in R} |F_{N,S}(x) - H_N(x)| = o(N^{-1/2}), \tag{1.17}$$

as $N \to \infty$.

As already indicated in our introduction the exisence of an Edgeworth expansion for $F_{N,S}$ was proved by Hall and Padmanabhan (1992). In (1.16) and (1.17) we give the precise and simple explicit form of an EE for $F_{N,S}$. In fact formally the form of our EE H_N (cf.(1.16)) coincides with the one given on p.1545 of Putter and van Zwet (1998). However, our Theorem 1.2 can not be inferred from the result of Putter and van Zwet (1998): the second condition in (1.18) is not satisfied for our T_N, that is, for a studentized trimmed mean (cf. Lemma A.2, Appendix).

REMARK 1.1. It is clear from the proofs of Theorems 1.1 and 1.2 that the order of the remainder term which we really obtain in relations (1.12) and (1.17) is $O((\log N)^{5/4}/N^{3/4})$, as $N \to \infty$.

To obtain empirical Edgeworth expansions (cf. Helmers (1991), Putter and van Zwet (1998)) we replace λ_1, λ_2, β_N and σ_W in (1.11) and (1.16) by statistical estimates. The estimation of λ_1 is straightforward. Let us define

$$\hat{\lambda}_1 = S_N^{-3} \hat{\gamma}_{3,W}$$
$$= S_N^{-3} \left(\frac{k}{N} (X_{k:N} - \hat{\mu}_W)^3 + \frac{1}{N} \sum_{i=k+1}^{m-1} (X_{i:N} - \hat{\mu}_W)^3 + \frac{N-m+1}{N} (X_{m:N} - \hat{\mu}_W)^3 \right)$$

($\hat{\mu}_W$ and S_N were defined in (1.13) and (1.14)) to be an estimate for λ_1. As to λ_2 and β_N, we first have to estimate the values of density $f(\xi_\alpha)$ and $f(\xi_\beta)$. We shall use kernel estimators with a simple step-like kernel. Put $g(x) = I_{\{|x| \le 1/2\}}$. Take the width of kernel $\Delta = N^{-1/4}$ and put

$g_\Delta(x) = \frac{1}{\Delta} g\left(\frac{x}{\Delta}\right) = \frac{1}{\Delta} I_{\{|x|\le \Delta/2\}}$, where $\int_{-\infty}^{\infty} g_\Delta(x)\,dx = 1$. Then our estimates for values of density at the quantiles where trimming occurs will be the following:

$$\hat{f}(\xi_\nu) = \frac{1}{N}\sum_{i=1}^{N} g_\Delta(X_i - X_{r:N}) = N^{-3/4}\sum_{i=1}^{N} I_{\{2N^{1/4}|X_i - X_{r:N}|\le 1\}}, \tag{1.18}$$

where $\nu = \alpha$ and $r = k$ or $\nu = \beta$ and $v = m$ respectively. Our estimates of $f(\xi_\alpha)$ and $f(\xi_\beta)$ are rather simple ones and sufficient for our purposes. (cf. also Reiss (1989), p.262 for related results) Thus, we obtain the following estimates for λ_2 and β_N:

$$\hat{\lambda}_2 = S_N^{-3}\left\{-\alpha^2(\hat{f}(\xi_\alpha))^{-1}[\hat{\mu}_W - X_{k:N}]^2 + (1-\beta)^2(\hat{f}(\xi_\beta))^{-1}[\hat{\mu}_W - X_{m:N}]^2\right\},$$

$$\hat{\beta}_N = \frac{1}{N}\left\{-(\alpha N - [\alpha N])\left(T_N - X_{k:N}\right) - \frac{1}{2}\alpha(1-\alpha)(\hat{f}(\xi_\alpha))^{-1}\right.$$
$$\left. +(\beta N - [\beta N])\left(T_N - X_{m:N}\right) + \frac{1}{2}\beta(1-\beta)(\hat{f}(\xi_\alpha))^{-1}\right\}.$$

When the conditions of Theorem 1.1 are satisfied, the estimates $\hat{\lambda}_1$, $\hat{\lambda}_2$ and $\hat{\beta}_N$ are consistent estimators of the corresponding quantities λ_1, λ_2 and β_N (cf. Sect.5). Replacing these latter quantities by their estimates in formulas (1.11) and (1.16), we obtain the empirical Edgeworth expansions:

$$\hat{G}_N(x) = \Phi(x) - \frac{\phi(x)}{6\sqrt{N}}\left((\hat{\lambda}_1 + 3\hat{\lambda}_2)(x^2-1) + 6N\frac{\hat{\beta}_N}{S_N}\right),$$

$$\hat{H}_N(x) = \Phi(x) + \frac{\phi(x)}{6\sqrt{N}}\left((2x^2+1)\hat{\lambda}_1 + 3(x^2+1)\hat{\lambda}_2 - 6N\frac{\hat{\beta}_N}{S_N}\right).$$

The result, establishing the validity of the empirical Edgeworth expansions, is given by the following assertion.

THEOREM 1.3. *Suppose that the conditions of Theorem 1.1 hold. Then*

$$\sup_{x\in R}|F_N(x) - \hat{G}_N(x)| = o_p\left(\frac{1}{\sqrt{N}}\right), \tag{1.19}$$

$$\sup_{x\in R}|F_{N,S}(x) - \hat{H}_N(x)| = o_p\left(\frac{1}{\sqrt{N}}\right). \tag{1.20}$$

as $N \to \infty$.

REMARK 1.2. It is clear from Remark 1.1 and the Lemma's 5.1 and 5.2 that we can strengthen (1.19) and (1.20) to $\sup_{x\in R}|F_N(x) - \hat{G}_N(x)| = O\left((\log N)^{5/4}N^{-3/4}\right)$ with probability $1 - O\left(N^{-c}\right)$, for every $c > 0$, as $N \to \infty$, and similarly, $\sup_{x\in R}|F_{N,S}(x) - \hat{H}_N(x)| = O\left((\log N)^{5/4}N^{-3/4}\right)$, except on a set with probability $O(N^{-c})$, for every $c > 0$.

2. AUXILIARY RESULTS

Let $U_1, \ldots, U_N$ are independent r.v.'s uniformly distributed on $(0,1)$ and $U_{1,N}, \ldots, U_{N,N}$ denote the corresponding order statistics. Define a binomial r.v.

$$N_\alpha = \sharp\{U_i : U_i \le \alpha\}.$$

The following lemma is a version of Bahadur's (1966) representation (cf. also Theorem 6.3.1, Reiss (1989)) for the sample quantile. In this section k is an integer, $k = \alpha N + O(1)$, $N \to \infty$.

LEMMA 2.1. *Suppose that $f = F'$ exists in neighborhood of ξ_α, satisfies a Lipschitz condition and $f(\xi_\alpha) > 0$. Let G be some function, having a derivative $g = G'$ in neighborhood of ξ_α, and suppose that g satisfies a Lipschitz condition in this neighborhood. Then*

$$G(X_{k:N}) = G(\xi_\alpha) - \frac{N_\alpha - \alpha N}{N} g(\xi_\alpha) \frac{1}{f(\xi_\alpha)} + R_N, \tag{2.1}$$

where

$$P(|R_N| > A(\log N/N)^{3/4}) = O(N^{-c}), \tag{2.2}$$

as $N \to \infty$, for every $c > 0$ and some $A > 0$, not depending on N.

We omit the proof because the lemma is essentially known and its proof requires similar arguments will also be used in the proof of Lemma 2.2. To state next lemma we shall adopt the following notation. Let $\sum_{i=k}^{m} (.)_i = sign[m-k] \sum_{i=k\wedge m}^{k \vee m} (.)_i$ for all integer k and m.

LEMMA 2.2. *Suppose that the conditions of lemma 2.1 are satisfied. Then*

$$\frac{1}{N} \sum_{i=k}^{N_\alpha} (G(X_{i:N}) - G(\xi_\alpha)) = -\frac{(N_\alpha - \alpha N)^2}{2N^2} g(\xi_\alpha) \frac{1}{f(\xi_\alpha)} + R_N, \tag{2.3}$$

where

$$P(|R_N| > A(\log N/N)^{5/4}) = O(N^{-c}), \tag{2.4}$$

as $N \to \infty$ for every $c > 0$ with some $A > 0$, not depending on N.

This lemma extends and sharpens the relations (3.2) and (3.3) given (for the case $G(x) = x$) in Hall and Padmanabhan (1992). Note also that the factor $(1-\alpha)^{-1}$ in formula (3.2) and $(1-\beta)^{-1}$ in formula (3.3) (see Hall and Padmanabhan (1992)) should be omitted.

COROLLARY 2.1. *Suppose that $f = F'$ exists in neighborhood of ξ_α, satisfies a Lipschitz condition and $f(\xi_\alpha) > 0$. Then*

$$\frac{1}{N} \sum_{i=k}^{N_\alpha} (X_{i:N} - \xi_\alpha) = -\frac{(N_\alpha - \alpha N)^2}{2N^2} \frac{1}{f(\xi_\alpha)} + R_{N,1}, \tag{2.5}$$

$$\frac{1}{N} \sum_{i=k}^{N_\alpha} \left(X_{i,N}^2 - \xi_\alpha^2\right) = -\frac{(N_\alpha - \alpha N)^2}{N^2} \xi_\alpha \frac{1}{f(\xi_\alpha)} + R_{N,2}, \tag{2.6}$$

where $R_{N,i}$, $i = 1, 2$, satisfy (2.4).

PROOF. We begin by writing

$$\frac{1}{N} \sum_{i=k}^{N_\alpha} (G(X_{i:N}) - G(\xi_\alpha)) \stackrel{\mathrm{d}}{=} \frac{1}{N} \sum_{i=k}^{N_\alpha} [(G \circ F^{-1})(U_{i:N}) - (G \circ F^{-1})(\alpha)]$$

$$= \frac{1}{N} g(\xi_\alpha) \frac{1}{f(\xi_\alpha)} \sum_{i=k}^{N_\alpha} (U_{i:N} - \alpha) + R_{N,3}, \tag{2.7}$$

where

$$|R_{N,3}| \le \frac{C}{N} \sum_{i=k\wedge N_\alpha}^{k\vee N_\alpha} (U_{i:N}-\alpha)^2 \le \frac{C|k-N_\alpha|}{N}\left[(U_{k:N}-\alpha)^2 \vee (U_{N_\alpha,N}-\alpha)^2\right] \tag{2.8}$$

with C equal to the product of Lipschitz's constants of functions g and h. Let us fix an arbitrary $c > 0$ and note that

$$P\left((\alpha - U_{N_\alpha,N})^2 > A_1 \log N/N\right) \le P\left(U_{N_\alpha+1,N} - U_{N_\alpha,N} > (A_1 \log N/N)^{1/2}\right) =$$

$$P\left(U_{1:N} > (A_1 \log N/N)^{1/2}\right) = O(N^{-c}). \tag{2.9}$$

Here and elsewhere A_j denote the positive constants which do not depend on N. Besides, by Bernstein's inequality

$$P(|N_\alpha - k| > (A_2 N \log N)^{1/2}) = O(N^{-c}), \tag{2.10}$$

with $A_2 = 2c_1\alpha(1-\alpha)$, and by lemma 3.1.1, Reiss (1989)

$$P((U_{k:N}-\alpha)^2 > A_3 \log N/N) = O(N^{-c}),$$

as $N \to \infty$. Therefore (2.8) implies that

$$P(|R_{N,3}| > A_4 (\log N/N)^{3/2}) = O(N^{-c}) \tag{2.11}$$

with $A_4 = CA_2 \max(A_1, A_3)$. Now consider the dominant term on the r.h.s. of (2.7). By (2.10) we can bound our quantities on the event $E = \left\{\omega : |N_\alpha - k| < (A_2 N \log N)^{1/2}\right\}$. Fix N and N_α for which the event E holds true. Without loss of generality let $k \le N_\alpha$. Note that conditional on N_α the order statistic $U_{i:N}$, $k \le i \le N_\alpha$, is distributed as i-th order statistic of the sample of size N_α from the uniform on $(0,\alpha)$ distribution and $E(U_{i:N}|N_\alpha) = \frac{\alpha i}{N_\alpha+1}$, for $i = k, \dots, N_\alpha$. Write

$$\frac{1}{N}\sum_{i=k}^{N_\alpha}(U_{i:N}-\alpha) = \frac{1}{N}\left[\sum_{i=k}^{N_\alpha}(U_{i:N} - \frac{\alpha i}{N_\alpha+1}) + \sum_{i=k}^{N_\alpha}(\frac{\alpha i}{N_\alpha+1} - \alpha)\right] =$$

$$\frac{1}{N}\sum_{i=k}^{N_\alpha}(U_{i:N} - \frac{\alpha i}{N_\alpha+1}) + \frac{\alpha}{(N_\alpha+1)N}\sum_{i=k}^{N_\alpha}(i - N_\alpha - 1) =$$

$$\frac{1}{N}\sum_{i=k}^{N_\alpha}(U_{i:N} - \frac{\alpha i}{N_\alpha+1}) - \frac{\alpha(N_\alpha - \alpha N)^2}{2NN_\alpha} + O(N^{-1}). \tag{2.12}$$

For the second term on the r.h.s. of (2.12) we have

$$-\frac{\alpha(N_\alpha-\alpha N)^2}{2NN_\alpha} = -\frac{(N_\alpha-\alpha N)^2}{2N}\frac{\alpha}{\alpha N + (N_\alpha - \alpha N)} = -\frac{(N_\alpha-\alpha N)^2}{2N^2} + R_{N,4}, \tag{2.13}$$

where in view of (2.10)

$$P(|R_{N,4}| > A_5 (\log N/N)^{3/2}) = O(N^{-c}) \tag{2.14}$$

as $N \to \infty$ with $A_5 = A_2^3$. For the first term on the r.h.s. of (2.12) we can write

$$\frac{1}{N}\left|\sum_{i=k}^{N_\alpha}\left(U_{i:N} - \frac{\alpha i}{N_\alpha+1}\right)\right| \le \frac{N_\alpha - k + 1}{N}\max_{k\le i\le N_\alpha}\left|U_{i:N} - \frac{\alpha i}{N_\alpha+1}\right|. \tag{2.15}$$

Note that we suppose that the event E holds true and (without loss of generality) that $k \le N_\alpha$ (otherwise a similar argument with respect to $(1-\alpha, 1)$ instead $(0,\alpha)$) will do). Fix an arbitrary$c_1 >$

$c+1/2$ and note that conditional on N_α the the variance of the order statistic $U_{i:N}$, $k \le i \le N_\alpha$, equals to $\frac{\alpha^2 i(N_\alpha - i+1)}{(N_\alpha+1)^2(N_\alpha+2)} = O\left((\log N)^{1/2} N^{-3/2}\right)$. By lemma 3.1.1, Reiss(1989), uniformly for $k \le i \le N_\alpha$

$$P\left(\left|U_{i:N} - \frac{\alpha i}{N_\alpha+1}\right| > A_6 (\log N/N)^{3/4} | N_\alpha\right) = O(N^{-c_1}), \tag{2.16}$$

as $N \to \infty$.Relations (2.15) and (2.16) together imply

$$P\left(\frac{1}{N}\left|\sum_{i=k}^{N_\alpha}(U_{i:N} - \frac{\alpha i}{N_\alpha+1})\right| > (A_2)^{1/2} A_6\ (\log N/N)^{5/4} \Big| N_\alpha\right) \le$$

$$(A_2 N \log N)^{1/2} O(N^{-c_1}) = O(N^{-c}), \tag{2.17}$$

as $N \to \infty$. Now (2.3) and (2.4) follow by (2.7), (2.10)–(2.14) and (2.17). The lemma is proved. □

3. Proof of Theorem 1.1.

To begin with let us note that we can replace T_N (cf. (1.1)) by

$$N^{-1/2} \sum_{i=k}^{m} X_{i:N}, \tag{3.1}$$

where $k = [\alpha N]+1$, $m = [\beta N]$, $0 < \alpha < \beta < 1$. Note that this will affect only the bias term (see Lemma A.1, Appendix), and we shall take that into account whenever needed. Define $I_\nu(X_i) = I_{\{X_i \le \xi_\nu\}}$, where $\xi_\nu = F^{-1}(\nu)$, $0 < \nu < 1$, and I_A is the indicator of event A. Then for the Winsorized r.v. W_i (cf. (1.2)) we can write

$$W_i = X_i I_\beta(X_i)(1 - I_\alpha(X_i)) + \xi_\alpha I_\alpha(X_i) + \xi_\beta(1 - I_\beta(X_i)). \tag{3.2}$$

Recall that μ_W, σ_W^2, $\gamma_{3,W}$ denote first three cumulants of r.v. EW_1 (cf.(1.3)). Define a $U-$statistic of degree 2 by

$$L_N + U_N = \sum_{i=1}^{N} L_{N,i} + \sum_{1 \le i} \sum_{< j \le N} U_{N,(i,j)}, \tag{3.3}$$

where

$$L_{N,i} = \frac{1}{\sqrt{N}}(W_i - \mu_W) =$$

$$\frac{1}{\sqrt{N}}\Big[X_i I_\beta(X_i)(1 - I_\alpha(X_i)) + \xi_\alpha I_\alpha(X_i) + \xi_\beta(1 - I_\beta(X_i)) - \mu_W\Big], \tag{3.4}$$

$$U_{N,(i,j)} = \frac{1}{N\sqrt{N}}\Big[- \frac{1}{f(\xi_\alpha)}(I_\alpha(X_i) - \alpha)(I_\alpha(X_j) - \alpha)$$

$$+ \frac{1}{f(\xi_\beta)}(I_\beta(X_i) - \beta)(I_\beta(X_j) - \beta)\Big]. \tag{3.5}$$

Note that

$$EL_{N,i} = 0 \tag{3.6}$$

for all $i = 1, \ldots, N$ and

$$EU_{N,(i,j)} = 0, \qquad E(L_{N,i}U_{N,(i,j)}) = 0 \tag{3.7}$$

for all $i, j = 1, \ldots, N$ $(i \ne j)$. Using (3.4)–(3.7), we easily check that

$$\sigma^2_{L_N+U_N} = E(L_N + U_N)^2 = E(L_N^2) + O(N^{-1}) = \sigma_W^2 + O(N^{-1}), \tag{3.8}$$

and also that

$$E(L_N+U_N)^3 = E(L_N^3) + 3E(L_N^2 U_N) + O(N^{-3/2}) =$$

$$\frac{1}{\sqrt{N}}\gamma_{3,W} + 3\frac{1}{\sqrt{N}}\left\{-\frac{1}{f(\xi_\alpha)}\left[E((W_1-\mu_W)(I_\alpha(X_1)-\alpha))\right]^2 + \right.$$

$$\left.\frac{1}{f(\xi_\beta)}\left[E((W_1-\mu_W)(I_\beta(X_1)-\beta))\right]^2\right\} + O(N^{-3/2}) =$$

$$\frac{1}{\sqrt{N}}\gamma_{3,W} + 3\frac{1}{\sqrt{N}}\left[-\frac{1}{f(\xi_\alpha)}\alpha^2\left[\xi_\alpha-\mu_W\right]^2 + \frac{1}{f(\xi_\beta)}(1-\beta)^2[\xi_\beta-\mu_W]^2\right] +$$

$$+O(N^{-3/2}). \tag{3.9}$$

Relations (3.8) and (3.9) imply that

$$E\left(\frac{L_N+U_N}{\sigma_{(L_N+U_N)}}\right)^3 = \frac{\lambda_1+3\lambda_2}{\sqrt{N}} + O(N^{-3/2}), \tag{3.10}$$

whereλ_1 and λ_2 as in (1.5).

The next lemma ensures that the approximation of T_N by a $U-$statistic of the form (3.3) has a remainder of classical Bahadur's order of magnitude $N^{-3/4}(\log N)^{5/4}$.

Lemma 3.1. *Suppose that the conditions of Theorem 1.1 hold. Then*

$$P\left(|T_N - ET_N' - (L_N+U_N)| > A(\log N)^{5/4}N^{-3/4}\right) = O(N^{-c}) \tag{3.11}$$

as $N\to\infty$, for every $c>0$ with some $A>0$ independent on N.

Proof of Lemma 3.1. Let $W_{i:N}$, $i=1,\dots,N$, denote the order statistics, corresponding to $W_1,\dots,W_N$. Put$N_\nu = \sharp\{X_i : X_i \le \xi_\nu\}$, $0<\nu<1$. Then

$$W_{i:N} = \begin{cases} \xi_\alpha\ , & i\le N_\alpha, \\ X_{i:N}\ , & N_\alpha < i \le N_\beta, \\ \xi_\beta\ , & i > N_\beta. \end{cases}$$

Now note that

$$T_N - \frac{1}{\sqrt{N}}\sum_{i=1}^N W_i = \frac{1}{\sqrt{N}}\left(\sum_{i=k}^m X_{i:N} - N_\alpha\xi_\alpha - \sum_{i=N_\alpha+1}^{N_\beta} X_{i:N} - (N-N_\beta)\xi_\beta\right) =$$

$$\frac{1}{\sqrt{N}}\left\{sign[N_\alpha-(k-1)]\sum_{i=k\wedge(N_\alpha+1)}^{N_\alpha\vee(k-1)} X_{i:N} - sign(N_\beta-m)\sum_{i=(m\wedge N_\beta)+1}^{m\vee N_\beta} X_{i:N}\right.$$

$$\left.-N_\alpha\xi_\alpha - (N-N_\beta)\xi_\beta\right\} = \frac{1}{\sqrt{N}}\left\{sign[N_\alpha-(k-1)]\sum_{i=k\wedge(N_\alpha+1)}^{N_\alpha\vee(k-1)}(X_{i:N}-\xi_\alpha)\right.$$

$$\left.-sign(N_\beta-m)\sum_{i=(m\wedge N_\beta)+1}^{m\vee N_\beta}(X_{i:N}-\xi_\beta) - (k-1)\xi_\alpha - (N-m)\xi_\beta\right\} =$$

$$-\frac{(N_\alpha-\alpha N)^2}{2N\sqrt{N}}\frac{1}{f(\xi_\alpha)} + \frac{(N_\beta-\beta N)^2}{2N\sqrt{N}}\frac{1}{f(\xi_\beta)} - \frac{k-1}{\sqrt{N}}\xi_\alpha - \frac{N-m}{\sqrt{N}}\xi_\beta + R_N,$$

where by Lemma 2.2

$$P\left(|R_N| > A(\log N)^{5/4}N^{-3/4}\right) = O(N^{-c}) \tag{3.12}$$

as $N \to \infty$, for every $c > 0$ with some $A > 0$ independent on N. Define

$$Q_N = -\frac{(N_\alpha - \alpha N)^2}{2N\sqrt{N}}\frac{1}{f(\xi_\alpha)} + \frac{(N_\beta - \beta N)^2}{2N\sqrt{N}}\frac{1}{f(\xi_\beta)} =$$

$$\frac{1}{2N\sqrt{N}}\left\{-\left[\sum_{i=1}^{N}(I_\alpha(X_i) - \alpha)\right]^2\frac{1}{f(\xi_\alpha)} + \left[\sum_{i=1}^{N}(I_\beta(X_i) - \beta)\right]^2\frac{1}{f(\xi_\beta)}\right\}$$

It is evident that Q_N is a symmetric polynomial of degree two with

$$E(Q_N) = \frac{1}{2\sqrt{N}}\left\{ -\alpha(1-\alpha)\frac{1}{f(\xi_\alpha)} + \beta(1-\beta)\frac{1}{f(\xi_\beta)}\right\}.$$

Note that

$$E\frac{1}{\sqrt{N}}\sum_{i=1}^{N}W_i = \sqrt{N}\mu_W = \sqrt{N}\left((\beta-\alpha)\mu(\alpha,\beta) + \alpha\xi_\alpha + (1-\beta)\xi_\beta\right). \tag{3.13}$$

Then we can write

$$\begin{aligned} T_N &= L_N + Q_N - EQ_N + \sqrt{N}\left((\beta-\alpha)\mu(\alpha,\beta) + \alpha\xi_\alpha + (1-\beta)\xi_\beta\right) \\ &-\frac{k-1}{\sqrt{N}}\xi_\alpha - \frac{N-m}{\sqrt{N}}\xi_\beta + \frac{1}{2\sqrt{N}}\left\{ -\alpha(1-\alpha)\frac{1}{f(\xi_\alpha)} + \beta(1-\beta)\frac{1}{f(\xi_\beta)}\right\} + R_N \\ &= L_N + Q_N - EQ_N + \sqrt{N}(\beta-\alpha)\mu(\alpha,\beta) + \frac{1}{\sqrt{N}}\Big\{ -[k-1-\alpha N]\xi_\alpha \\ &-\frac{1}{2}\frac{1}{f(\xi_\alpha)}\alpha(1-\alpha) + [m-\beta N]\xi_\beta + \frac{1}{2}\frac{1}{f(\xi_\beta)}\beta(1-\beta)\Big\} + R_N. \end{aligned} \tag{3.14}$$

Let us compare the expression within curly brackets on the r.h.s. of (3.14) with the formula for B_2 (a bias term for T_N,given by (3.1)) (cf.(A.3), Appendix). As a result we obtain the next formulas:

$$T_N - \sqrt{N}(\beta-\alpha)\mu(\alpha,\beta) = L_N + Q_N - EQ_N + \sqrt{N}B_2 + R_N, \tag{3.15}$$

$$T_N - ET'_N = L_N + Q_N - EQ_N + R_N, \tag{3.16}$$

and since $ET'_N - (\beta-\alpha)\mu(\alpha,\beta) = \sqrt{N}B_2 + O(^{-3/2})$ (cf. (A.3),Appendix), R_N satisfies (3.15). For the quantity $Q_N - EQ_N$ we can write

$$Q_N - EQ_N = U_N + \frac{\bar{r}_N}{2\sqrt{N}}, \tag{3.17}$$

where U_N as in (3.3) and

$$\begin{aligned} \bar{r}_N = \frac{1}{N}\sum_{i=1}^{N}\Big\{ &-\frac{1}{f(\xi_\alpha)}[(I_\alpha(X_i)-\alpha)^2 - \alpha(1-\alpha)] \\ &+\frac{1}{f(\xi_\beta)}[(I_\beta(X_i)-\beta)^2 - \beta(1-\beta)]\Big\}. \end{aligned}$$

Note that $\bar{r}_N$ is the average of N i.i.d. bounded and centered ($E\bar{r}_N = 0$) r.v.'s, and by Hoeffding's inequality

$$P\left(|\bar{r}_N| > A(\log N/N)^{1/2}\right) = O(N^{-c}) \tag{3.18}$$

for every $c > 0$ and some $A > 0$, not depending on N. Therefore $\frac{1}{2}\bar{r}_N/\sqrt{N}$ on the r.h.s. of (3.17) is negligible for our purposes. Relations (3.12) and (3.16)–(3.18) together imply (3.11). The lemma is proved. □

Proof of Theorem 1.1. Using the Lemma 3.1 and Lemma A.1 (cf.Appendix), for *df* of T_N (cf.(1.1)) defined by (1.10) we can write

$$F_{T_N}(x) = P\left\{\frac{N^{1/2}(\beta-\alpha)(T_N - ET'_N)}{\sigma_W} \le x - \frac{N^{1/2}\beta_N + O(N^{-1})}{\sigma_W}\right\} =$$

$$P\left\{\frac{L_N + U_N}{\sigma_W} \le \frac{[\beta N] - [\alpha N]}{(\beta-\alpha)N}\left(x - \frac{N^{1/2}\beta_N}{\sigma_W} + O(N^{-1})\right) - \frac{R_N}{\sigma_W}\right\} =$$

$$P\left\{\frac{L_N + U_N}{\sigma_W} \le x(1 + O(N^{-1})) - \frac{N^{1/2}\beta_N}{\sigma_W} - \frac{R_N}{\sigma_W} + O(N^{-1})\right\}, \tag{3.19}$$

where $L_N + U_N$ (cf.(3.3)) is U-statistic of degree two with the canonical functions

$$g_N(x) = E(L_N + U_N | X_1 = x) =$$

$$\frac{1}{\sqrt{N}}[xI_\beta(x)(1 - I_\alpha(x)) + \xi_\alpha I_\alpha(x) + \xi_\beta(1 - I_\beta(x)) - \mu_W],$$

$$\psi_N(x,y) = E(L_N + U_N | X_1 = x, X_2 = y) - g_N(x) - g_N(y) =$$

$$\frac{1}{N\sqrt{N}}\left[-(I_\alpha(x) - \alpha)(I_\alpha(y) - \alpha)\frac{1}{f(\xi_\alpha)} + (I_\beta(x) - \beta)(I_\beta(y) - \beta)\frac{1}{f(\xi_\beta)}\right],$$

where

$$E(g_N(X_1)) = 0, \qquad E(\psi_N(X_1, X_2)) = 0,$$

$$E(\psi_N(X_1, X_2)|X_2) = 0 \quad a.s.$$

The local smoothness assumption of our theorem directly yields that the distribution of r.v. $g_N(X_1) = \frac{1}{\sqrt{N}}(W_1 - \mu_W)$ has a nontrivial absolutely continuous component and Cramer's condition

$$(C) \qquad \limsup_{|t|\to\infty} |E\exp\{it\sqrt{N}g_N(X_1)\}| < \infty$$

is satisfied. Since the functions $\sqrt{N}g_N(x)$ and $N^{3/2}\psi_N(x,y)$ are both bounded, we trivially have that

$$\beta_4 = E\left(\sqrt{N}g_N(X_1)\right)^4 < \infty,$$

$$\gamma_3 = E\left|N^{3/2}\psi_N(X_1, X_2)\right|^3 < \infty.$$

Therefore, we can apply Thm.1.2 of Bentkus, Götze and van Zwet (1997) (note that the quantity Δ_3^2 appearing in Thm.1.2 of Bentkus et. all (1997) is zero in our case). Define $F_N(x) = \Phi(x) - \phi(x)\frac{\lambda_1 + 3\lambda_2}{6\sqrt{N}}(x^2 - 1)$, where λ_1 and λ_2 as in (1.5) (cf.also (3.10)). Then by Thm.1.2 (Bentkus et.all (1997))

$$\sup_{x\in R}\left|P\left\{\frac{L_N + U_N}{\sigma_W} \le x\right\} - F_N(x)\right| = O(N^{-1}).$$

For R_N we have the bound (3.12), that is $|R_N| = O((\log N)^{5/4}N^{-3/4})$ with probability $1 - o(N^{-c})$ for every $c > 0$. Therefore, as $F_N'(x)$ and $xF_N'(x)$ are the bounded functions, we obtain on the r.h.s. of (3.19)

$$F_N(x) - \frac{\sqrt{N}\beta_N}{\sigma_W} + O((\log N)^{5/4}N^{-3/4}) =$$
$$G_N(x) + O((\log N)^{5/4}N^{-3/4}).$$

This proves (1.12) and Theorem 1.1. □

4. PROOF OF THEOREM 1.2.

Let S_N^2 be (cf.(1.14)) the plug-in estimator for σ_W^2 (cf.(1.3)). The following lemma is a modification of Lemma 4.3 of Putter and van Zwet (1998), appropriate for our purposes.

LEMMA 4.1. *Suppose that the assumptions of Theorem 1.1 hold. Then*

$$P\left(|S_N^2 - \sigma_W^2 - V_N| > A(\log N/N)^{3/4}\right) = O(N^{-c}) \tag{4.1}$$

as $N \to \infty$ for every $c > 0$ and some $A > 0$, not depending on N, where

$$V_N = V_{N,1} + V_{N,2}\ , \tag{4.2}$$

$$V_{N,1} = 2\alpha\frac{1}{f(\xi_\alpha)}\frac{N_\alpha - \alpha N}{N}[\mu_W - \xi_\alpha] + 2(1-\beta)\frac{1}{f(\xi_\beta)}\frac{N_\beta - \beta N}{N}[\mu_W - \xi_\beta],$$

$$V_{N,2} = \frac{1}{N}\sum_{i=1}^{N}\left[(W_i - \mu_W)^2 - \sigma_W^2\right]\ .$$

Moreover,

$$E(V_N) = 0\ ;\quad E(V_N^2) = O(N^{-1}) \tag{4.3}$$

as $N \to \infty$.

This lemma essentially asserts that the difference between σ_W^2 and its estimator S_N^2 can be expressed as a sum of i.i.d. r.v.'s plus a remainder term which is of negligible order for our purposes.

PROOF. Define the auxiliary quantity

$$S_W^2 = \frac{1}{N}\sum_{i=1}^{N} W_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} W_i\right)^2 =$$

$$\frac{N_\alpha}{N}\xi_\alpha^2 + \frac{1}{N}\sum_{i=N_\alpha+1}^{N_\beta} X_{i:N}^2 + \frac{N-N_\beta}{N}\xi_\beta^2\ - \left(\frac{N_\alpha}{N}\xi_\alpha + \frac{1}{N}\sum_{i=N_\alpha+1}^{N_\beta} X_{i:N} + \frac{N-N_\beta}{N}\xi_\beta\right)^2.$$

First we prove that

$$S_N^2 = S_W^2 + V_{N,1} + R_{N,1}, \tag{4.4}$$

Here and elsewhere $R_{N,1}$, $R_{N,1}^{(r)}$, $r = 1, 2, \ldots$ denote the remainder terms of Bahadur's order, satisfying (2.2). We have

$$S_N^2 - S_W^2 = \left[\frac{k}{N}X_{k:N}^2 + \frac{1}{N}\sum_{i=k+1}^{m} X_{i:N}^2 + \frac{N-m}{N}X_{m:N}^2 - \right.$$

$$\frac{N_\alpha}{N}\xi_\alpha^2 - \frac{1}{N}\sum_{i=N_\alpha+1}^{N_\beta} X_{i:N}^2 - \frac{N-N_\beta}{N}\xi_\beta^2\Bigg] + \left[\left(\frac{1}{N}\sum_{i=1}^{N} W_i\right)^2 - \hat{\mu}_W^2\right]. \tag{4.6}$$

Rewrite the term within the first square brackets on the r.h.s. of (4.6) as

$$\frac{k}{N}(X_{k:N}^2-\xi_\alpha^2)+\frac{1}{N}\sum_{i=k+1}^{N_\alpha}(X_{i:N}^2-\xi_\alpha^2)+\frac{N-m}{N}(X_{m:N}^2-\xi_\beta^2)-\frac{1}{N}\sum_{i=m+1}^{N_\beta}(X_{i:N}^2-\xi_\beta^2)$$

(cf. the proof of Lemma 3.1, above). By Lemmas 2.1 and 2.2 this equals to

$$-2\alpha\xi_\alpha\frac{1}{f(\xi_\alpha)}\frac{N_\alpha-\alpha N}{N}-\frac{(N_\alpha-\alpha N)^2}{N^2}\xi_\alpha\frac{1}{f(\xi_\alpha)}+R_{N,1}^{(1)}-$$

$$2(1-\beta)\xi_\beta\frac{1}{f(\xi_\beta)}\frac{N_\beta-\beta N}{N}+\frac{(N_\beta-\beta N)^2}{N^2}\xi_\beta\frac{1}{f(\xi_\beta)}+R_{N,1}^{(2)},$$

and by Bernstein's inequality for binomial r.v.'s N_α and N_β the latter expression reduces to

$$-2\alpha\xi_\alpha\frac{1}{f(\xi_\alpha)}\frac{N_\alpha-\alpha N}{N}-2(1-\beta)\xi_\beta\frac{1}{f(\xi_\beta)}\frac{N_\beta-\beta N}{N}+R_{N,1}^{(3)}. \tag{4.7}$$

Now consider the term within the second square brackets at the r.h.s. of (4.6). Argueing as before, we can rewrite this expression as

$$\left(\frac{2}{N}\sum_{i=1}^{N}W_i-\alpha\frac{1}{f(\xi_\alpha)}\frac{N_\alpha-\alpha N}{N}-(1-\beta)\frac{1}{f(\xi_\beta)}\frac{N_\beta-\beta N}{N}+R_{N,1}^{(4)}\right)\cdot$$

$$\left(\alpha\frac{1}{f(\xi_\alpha)}\frac{N_\alpha-\alpha N}{N}+(1-\beta)\frac{1}{f(\xi_\beta)}\frac{N_\beta-\beta N}{N}+R_{N,1}^{(5)}\right)=$$

$$\frac{2}{N}\left(\sum_{i=1}^{N}W_i\right)\left(\alpha\frac{1}{f(\xi_\alpha)}\frac{N_\alpha-\alpha N}{N}+(1-\beta)\frac{1}{f(\xi_\beta)}\frac{N_\beta-\beta N}{N}\right)+R_{N,1}^{(6)}\ . \tag{4.8}$$

The relations (4.6)–(4.8) together imply that

$$S_N^2-S_W^2=V_{N,1}+R_N+R_{N,1}^{(7)}, \tag{4.9}$$

where

$$R_N=2\left[\alpha\frac{1}{f(\xi_\alpha)}\frac{N_\alpha-\alpha N}{N}+(1-\beta)\frac{1}{f(\xi_\beta)}\frac{N_\beta-\beta N}{N}\right]\frac{1}{N}\sum_{i=1}^{N}(W_i-\mu_W)\ .$$

Note that W_i, $i=1,\ldots,N$, are bounded i.i.d. *r.v.*'s. Therefore by Hoeffding's inequality $\frac{1}{N}\left|\sum_{i=1}^N(W_i-\mu_W)\right| = O\left((\log N/N)^{1/2}\right)$ as $N\to\infty$ with probability $1-o(N^{-c})$ for every $c>0$. Combining the latter bound with Bernstein's inequality for the binomial r.v.'s N_α and N_β, we obtain that $|R_N|=O(\log N/N)$ with probability $1-o(N^{-c})$ for every $c>0$. Therefore (4.9) implies (4.4).

Next we prove that

$$S_W^2=\sigma_W^2+V_{N,2}+R_{N,2}, \tag{4.10}$$

where $|R_{N,2}|=O(\log N/N)$ with probability $1-o(N^{-c})$ for every $c>0$. We have

$$S_W^2-\sigma_W^2-V_{N,2}=S_W^2-\frac{1}{N}\sum_{i=1}^{N}(W_i-\mu_W)^2=-(\overline{W}-\mu_W)^2=R_{N,2}.$$

The application of Hoeffding's inequality to the bounded i.i.d. *r.v.*'s W_i proves (4.10). Relations (4.4) and (4.10) together imply (4.1). The lemma is proved. □

Now we turn to the proof of our result concerning the Studentized version of trimmed mean.

Proof of Theorem 1.2. Our proof of this theorem closely resembles the proof of Theorem 1.2 of Putter and van Zwet (1998). For the *df* $F_{N,S}(x)$ (cf.(1.15)) of a studentized trimmed mean we have

$$F_{N,S}(x) = P\left\{\frac{L_N+U_N}{S_N} \le \left(1+O(N^{-1})\right)\left[x - \frac{N^{1/2}\beta_N + O(N^{-1})}{S_N}\right] + \frac{R_{N,1}}{S_N}\right\} \tag{4.11}$$

(cf.(3.19)). Here and elsewhere $R_{N,1}$ denotes a remainder, which satisfies (3.12) and which can be different from line to line. Lemma 4.1 and Hoeffding's inequality for r.v. V_N together imply that $\left|\frac{1}{S_N} - \frac{1}{\sigma_W}\right| = O((\log N/N)^{1/2})$ with probability $1 - O(N^{-c})$ as $N \to \infty$ for every $c > 0$ (cf.also Lemma 5.2, below). Therefore, the r.h.s. of (4.11) equals to

$$P\left\{\frac{L_N+U_N}{S_N} \le \left(1+O(N^{-1})\right)\left[x - \frac{\sqrt{N}\beta_N}{\sigma_W}\right] + R_{N,1}\right\}. \tag{4.12}$$

Our aim now is to prove that

$$\sup_{x\in R}|F_{N,S}(x) - H_N(x)| = O\left((\log N)^{5/4}/N^{3/4}\right) \tag{4.13}$$

as $N \to \infty$ (this implies (1.17)). Define $\tilde{H}_N(x) = H_N(x) + \sigma_W^{-1}\sqrt{N}\beta_N\phi(x)$ (i.e. $\tilde{H}_N(x)$ is $H_N(x)$ without bias term). Since $H'_N(x)$ and $xH'_N(x)$ are bounded, relations (4.11) and (4.12) imply that it is sufficient to show that

$$\sup_{x\in R}|F_{(L_N+U_N)/S_N}(x) - \tilde{H}_N(x)| = O\left((\log N)^{5/4}/N^{3/4}\right), \tag{4.14}$$

where $F_{(L_N+U_N)/S_N}(x) = P\left((L_N+U_N)/S_N \le x\right)$. The application of the Lemma 4.1 yields that

$$F_{(L_N+U_N)/S_N}(x) = P\left(\frac{L_N+U_N}{\sigma_W} \le x\frac{(\sigma_W^2 + V_N + R_N)^{1/2}}{\sigma_W}\right),$$

where R_N is a remainder of Bahadur's order (i.e. satisfying (2.2)). Since $x\tilde{H}'_N(x)$ is bounded, it is sufficient to prove (4.14) with $F_{(L_N+U_N)/S_N}(x)$ replaced by

$$P\left(\frac{L_N+U_N}{\sigma_W} \le x\frac{(\sigma_W^2+V_N)^{1/2}}{\sigma_W}\right) = P\left(\frac{L_N+U_N}{\sigma_W} - x\left\{\left(1+\frac{V_N}{\sigma_W^2}\right)^{1/2} - 1\right\} \le x\right).$$

Following to Putter and van Zwet (1998), we also use inequality $1+\frac{z}{2}-\frac{z^2}{4} \le (1+z)^{1/2} \le 1+\frac{z}{2}$ ($|z| \le \frac{4}{5}$) to find that $\frac{V_N}{2\sigma_W^2} - \frac{V_N^2}{4\sigma_W^4} \le \left(1+\frac{V_N}{\sigma_W^2}\right)^{1/2} - 1 \le \frac{V_N}{2\sigma_W^2}$ (with probability $1-O(N^{-c})$, $c>0$). Since by Hoeffding's inequality $V_N^2 = O(\log N/N)$ with probability $1-O(N^{-c})$ for every $c>0$, we can replace $F_{(L_N+U_N)/S_N}(x)$ in (4.14) by $P\left(\frac{L_N+U_N}{\sigma_W} - x\frac{V_N}{2\sigma_W^2} \le x\right)$. Now it remains to show that

$$\sup_{x\in R}\left|P\left(\frac{L_N+U_N}{\sigma_W} - x\frac{V_N}{2\sigma_W^2} \le x\right) - \tilde{H}_N(x)\right| = O\left((\log N)^{5/4}/N^{3/4}\right), \tag{4.15}$$

as $N \to \infty$. First we prove (4.15), taking supremum for $x : |x| < \log N$ (cf. Putter and van Zwet (1998)). Note that $U_{Nx} = \frac{L_N+U_N}{\sigma_W} - x\frac{V_N}{2\sigma_W^2}$ is centered U-statistic of degree two with bounded (uniformly for all x: $|x| < \log N$) kernel. Moreover, U_{Nx} has nontrivial absolutely continuous component and Cramer's condition is satisfied. Theorem 1.1 of Bentkus, Götze and van Zwet (1997) now yields that

$$\sup_{|x|<\log N} \left| P\left(\frac{L_N + U_N}{\sigma_W} - \frac{xV_N}{2\sigma_W^2} \le x \right) - \tilde{G}_N(x) \right| = O\left(N^{-1}\right), \tag{4.16}$$

where $\tilde{G}_N(x) = \Phi\left(\frac{x}{\sigma_x}\right) - \frac{k_{3x}}{6\sigma_x^3}\left[\left(\frac{x}{\sigma_x}\right)^2 - 1\right]\phi\left(\frac{x}{\sigma_x}\right)$, $\sigma_x^2 = Var(U_{Nx}) = E\left(\frac{L_N+U_N}{\sigma_W} - \frac{xV_N}{2\sigma_W^2}\right)^2$ and $k_{3x} = E\left(\frac{L_N+U_N}{\sigma_W} - \frac{xV_N}{2\sigma_W^2}\right)^3$. Using formulas (3.3)–(3.5) and relations (4.2)–(4.3), we find that $\sigma_x^2 = 1 + O\left(\frac{\log N}{\sqrt{N}}\right)$ and $k_{3x} = \frac{\lambda_1 + 3\lambda_2}{\sqrt{N}} + O\left(\frac{\log N}{N}\right)$. Therefore

$$\tilde{G}_N(x) = \Phi\left(\frac{x}{\sigma_x}\right) - \frac{\lambda_1 + 3\lambda_2}{6\sqrt{N}}(x^2 - 1)\phi(x) + O\left(\frac{\log N}{N}\right) \tag{4.17}$$

(for $|x| < \log N$), that is σ_x influences the form of EE only through the term $\Phi\left(\frac{x}{\sigma_x}\right)$ (cf. Putter and van Zwet (1998)). For σ_x^2 we can write $\sigma_x^2 = E\left(\frac{L_N+U_N}{\sigma_W} - \frac{xV_N}{2\sigma_W^2}\right)^2 = 1 - x\sigma_W^{-3}E[(L_N+U_N)V_N] + O\left(\frac{\log^2 N}{N}\right)$. As U_N and V_N are uncorrelated, using formulas (3.3)–(3.4) and (4.2), we can write $E[(L_N+U_N)V_N] = E(L_N V_N) = \frac{1}{\sqrt{N}}(\gamma_{3,W} + 2\delta_{2,W})$. Thus, we obtain that $\sigma_x^2 = 1 - \frac{x(\lambda_1+2\lambda_2)}{\sqrt{N}} + O\left(\frac{\log^2 N}{N}\right)$ (cf. notations (1.3)–(1.5)). This implies that

$$\Phi\left(\frac{x}{\sigma_x}\right) = \Phi(x) + \phi(x)\frac{1}{2}\frac{x^2(\lambda_1 + 2\lambda_2)}{\sqrt{N}} + O\left(\frac{\log^2 N}{N}\right). \tag{4.18}$$

Relations (4.17) and (4.18) together yield that $\tilde{G}_N(x) = \tilde{H}_N(x) + O\left(\frac{\log^2 N}{N}\right)$ for $|x| < \log N$. To treat the case $|x| \ge \log N$, we use the same arguments as on p.1561 of Putter and van Zwet (1998). Thus, $\sup_{x\in R}\left| P\left(\frac{L_N+U_N}{\sigma_W} - \frac{xV_N}{2\sigma_W^2} \le x\right) - \tilde{H}_N(x)\right| = O\left(\frac{\log^2 N}{N}\right)$. This proves (4.15) and the theorem. □

5. Proof of Theorem 1.3.

In this section we state and prove lemmas on the (rate of) consistency of the estimators for λ_1, λ_2 and β_N. The validity of Theorem 1.3 follows directly from Theorem 1.1, 1.2 and these lemmas. In the first lemma we obtain the rate of convergence for our kernel estimates of the density evaluated at given quantiles, defined by (1.18).

Lemma 5.1. *Suppose that $f = F'$ exists in a neighborhood of ξ_α, satisfies a Lipschitz condition and $f(\xi_\alpha) > 0$. Then*

$$P\left(|\hat{f}(\xi_\alpha) - f(\xi_\alpha)| > A(\log N)^{1/2}/N^{1/4}\right) = O(N^{-c}) \tag{5.1}$$

as $N \to \infty$, for every $c > 0$ and some $A > 0$, not depending on N.

Proof. Define two quantities

$$\nu_{k,N} = \sharp\left\{X_i : |X_i - X_{k:N}| \le N^{-1/4}/2\right\}, \quad \nu_{\alpha,N} = \sharp\left\{X_i : |X_i - \xi_\alpha| \le N^{-1/4}/2\right\}. \tag{5.2}$$

Note that $E\nu_{\alpha,N} = N\int_{\xi_\alpha - N^{-1/4}/2}^{\xi_\alpha + N^{-1/4}/2} f(x)\,dx$, therefore we can write

$$\hat{f}(\xi_\alpha) - f(\xi_\alpha) = N^{-3/4}\nu_{k,N} - f(\xi_\alpha) =$$

$$N^{-3/4}\nu_{\alpha,N} + N^{-3/4}(\nu_{k,N} - \nu_{\alpha,N}) - f(\xi_\alpha) = Q_{1,N} + Q_{2,N} + Q_{3,N}, \tag{5.3}$$

where

$$Q_{1,N} = N^{-3/4}(\nu_{\alpha,N} - E\nu_{\alpha,N}), \qquad Q_{2,N} = N^{-3/4}(\nu_{k,N} - \nu_{\alpha,N}),$$

$$Q_{3,N} = N^{1/4} \int_{\xi_\alpha - N^{-1/4}/2}^{\xi_\alpha + N^{-1/4}/2} (f(x) - f(\xi_\alpha)\, dx.$$

For $Q_{1,N}$ we can write $Q_{1,N} = N^{1/4}\left(\bar{\nu}_{\alpha,N} - E\bar{\nu}_{\alpha,N}\right)$, where $\bar{\nu}_{\alpha,N} = \frac{1}{N}\sum_{i=1}^{N} I_{\{2N^{1/4}|X_i - \xi_\alpha| \le 1\}}$ is a mean of i.i.d. bounded r.v.'s. Therefore, by Hoeffding's inequality

$$P\left(|Q_{1,N}| > A_1(\log N)^{1/2}/N^{1/4}\right) = O(N^{-c}) \tag{5.4}$$

for every $c > 0$, as $N \to \infty$. Here and elsewhere A_i, $i = 1, 2, \ldots$ denote positive constants, not depending on N. Since $P(|X_{k:N} - \xi_\alpha| > A_2(\log N/N)^{1/2}) = O(N^{-c})$, for $Q_{2,N}$ we have with probability $1 - O(N^{-c})$

$$|Q_{2,N}| \le N^{-3/4}(\nu_{l,N} + \nu_{r,N}), \tag{5.5}$$

where $\nu_{l,N} = \sharp\left\{X_i : |X_i - \xi_\alpha + N^{-1/4}/2| \le A_2(\log N/N)^{1/2}\right\}$, $\nu_{r,N} = \sharp\left\{X_i : |X_i - \xi_\alpha - N^{-1/4}/2| \le A_2(\log N/N)^{1/2}\right\}$. Since $(\nu_{l,N} + \nu_{r,N})$ is a Binomial r.v. with parameter $p_N = O\left((\log N/N)^{1/2}\right)$ and $E(\nu_{l,N} + \nu_{r,N}) = O\left(N^{1/2}(\log N)^{1/2}\right)$, $\sigma_{\nu_{l,N}+\nu_{r,N}} = O\left(N^{1/4}(\log N)^{1/4}\right)$, by Bernstein inequality, with probability $1 - O(N^{-c})$, we have the following bound

$$|Q_{2,N}| \le A_3 N^{-1/4}(\log N)^{1/2}. \tag{5.6}$$

Finally for $Q_{3,N}$ the Lipschitz condition directly yields that

$$|Q_{3,N}| \le CN^{1/4} \int_{\xi_\alpha - N^{-1/4}/2}^{\xi_\alpha + N^{-1/4}/2} |x - \xi_\alpha|\, dx = \frac{1}{4} CN^{-1/4}, \tag{5.7}$$

where C is the Lipschitz's constant. Relations (5.3)–(5.7) imply (5.1). The lemma is proved.□

Let $\mu_{r,W} = EW_i^r = \int_0^1 Q^r(u)\, du$ denotes the r-th moment of W_i for any positive integer r and let $\hat{\mu}_{r,W} = \frac{k}{N} X_{k:N}^r + \frac{1}{N}\sum_{i=k+1}^{m} X_{i:N}^r + \frac{N-m}{N} X_{m:N}^r$ be the plug-in estimator for $\mu_{r,W}$.

LEMMA 5.2. *Suppose that $f = F'$ exists in neighborhoods of the points ξ_α and ξ_β, satisfies a Lipschitz condition and $f(\xi_\nu) > 0$, $\nu = \alpha, \beta$. Then*

$$P\left(|\hat{\mu}_{r,W} - \mu_{r,W}| > A(\log N/N)^{1/2}\right) = O(N^{-c}) \tag{5.8}$$

as $N \to \infty$ for every $c > 0$ with some $A > 0$, not depending on N.

PROOF. Put $\bar{W}_r = \frac{1}{N}\sum_{i=1}^{N} W_i^r$, where W_i defined by (1.2), and note that as well as in (3.4) we can write

$$\bar{W}_r = \frac{N_\alpha}{N}\xi_\alpha^r + \frac{1}{N}\sum_{i=N_\alpha+1}^{N_\beta} X_{i:N}^r + \frac{N - N_\beta}{N}\xi_\beta^r.$$

We have

$$\hat{\mu}_{r,W} - \mu_{r,W} = (\hat{\mu}_{r,W} - \bar{W}_r) + (\bar{W}_r - \mu_{r,W}). \tag{5.9}$$

Note that $E\bar{W}_r = \mu_{r,W}$, therefore by Hoeffding inequality for the average of i.i.d. bounded r.v.'s $E\bar{W}_r$ we have $|\bar{W}_r - \mu_{r,W}| = O\left((\log N/N)^{1/2}\right)$ with probability $1 - O(N^{-c})$ for every $c > 0$. For $(\hat{\mu}_{r,W} - \bar{W}_r)$ on the r.h.s. of (5.9) we have

$$\hat{\mu}_{r,W} - \bar{W}_r = \frac{k}{N}(X_{k:N}^r - \xi_\alpha^r) + \frac{1}{N}\sum_{i=k+1}^{N_\alpha}(X_{i:N}^r - \xi_\alpha^r) +$$

$$\frac{N-m}{N}(X_{m:N}^r-\xi_\beta^r)-\frac{1}{N}\sum_{i=m+1}^{N_\beta}(X_{i:N}^r-\xi_\beta^r).$$

(cf.(4.6)). By Lemmas 2.1 and 2.2 the last expression equals to

$$-\alpha r\xi_\alpha^{r-1}\frac{N_\alpha-\alpha N}{N}\frac{1}{f(\xi_\alpha)}-\frac{(N_\alpha-\alpha N)^2}{2N^2}r\xi_\alpha^{r-1}\frac{1}{f(\xi_\alpha)}$$

$$-(1-\beta)r\xi_\beta^{r-1}\frac{N_\beta-\beta N}{N}\frac{1}{f(\xi_\beta)}+\frac{(N_\beta-\beta N)^2}{2N^2}r\xi_\beta^{r-1}\frac{1}{f(\xi_\beta)}+R_N, \tag{5.10}$$

where R_N is remaider term of the Bahadur's order (cf. (2.2)). Thus, by Bernstein inequality we find that

$$|\hat{\mu}_{r,W}-\bar{W}_r|=O\left((\log N/N)^{1/2}\right) \tag{5.11}$$

with probability $1-O(N^{-c})$ for every $c>0$. Relations (5.9)–(5.11) together imply (5.8). The lemma is proved. □

APPENDIX

In this appendix we first establish an asymptotic approximation for bias of T_N' (cf. (1.9)) in estimating of $\mu(\alpha,\beta)$. Secondly we prove that our Theorem 1.2 can not be inferred from Theorem 1.2 of Putter and van Zwet (1998) for studentized symmetric statistics..

LEMMA A.1. *Suppose the conditions of Theorem 1.1 are satisfied. Then*

$$b_N=\beta_N+O(N^{-3/2}), \tag{A.1}$$

whith b_N and β_N as in (1.8) and (1.9).

PROOF. To begin with we note that b_N (cf. (1.9) can be written as B_1+B_2 where $B_1=(\beta-\alpha)ET_N'-E\left(\frac{1}{N}\sum_{i=[\alpha N]+1}^{[\beta N]}X_{i:N}'\right)$ and $B_2=E\left(\frac{1}{N}\sum_{i=[\alpha N]+1}^{[\beta N]}X_{i:N}'\right)-(\beta-\alpha)\mu(\alpha,\beta)$. First we consider B_2. By a simple conditioning argument we have that B_2 equals (with $k=[\alpha N]+1$, $m=[\beta N]$)

$$\frac{1}{N}E\left(F^{-1}(U_{k:N})+F^{-1}(U_{m:N})+(m-k-1)\frac{\int_{U_{k:N}}^{U_{m:N}}F^{-1}(u)\,du}{U_{m:N}-U_{k:N}}\right)-(\beta-\alpha)\mu(\alpha,\beta). \tag{A.2}$$

Define

$$I(v_1,v_2)=\frac{\int_{v_1}^{v_2}F^{-1}(u)\,du}{v_1-v_2},\qquad I(\alpha,\beta)=\mu(\alpha,\beta).$$

The first and second partial derivatives are given by

$$\left.\frac{\partial I}{\partial v_1}\right|_{(\alpha,\beta)}=\frac{-\xi_\alpha+\mu(\alpha,\beta)}{\beta-\alpha},\qquad \left.\frac{\partial I}{\partial v_2}\right|_{(\alpha,\beta)}=\frac{\xi_\beta-\mu(\alpha,\beta)}{\beta-\alpha},$$

$$\left.\frac{\partial^2 I}{\partial v_1^2}\right|_{(\alpha,\beta)}=-\frac{2}{\beta-\alpha}\left[\frac{1}{2f(\xi_\alpha)}-\frac{\mu(\alpha,\beta)-\xi_\alpha}{\beta-\alpha}\right],$$

$$\left.\frac{\partial^2 I}{\partial v_2^2}\right|_{(\alpha,\beta)}=\frac{2}{\beta-\alpha}\left[\frac{1}{2f(\xi_\beta)}+\frac{\mu(\alpha,\beta)-\xi_\beta}{\beta-\alpha}\right],$$

$$\left.\frac{\partial^2 I}{\partial v_1\partial v_2}\right|_{(\alpha,\beta)}=\frac{\xi_\alpha+\xi_\beta-2\mu(\alpha,\beta)}{(\beta-\alpha)^2}.$$

A Taylor expansion argument now yields that (A.2) reduces to

$$\frac{1}{N}E\Big(F^{-1}(U_{k:N})+\ F^{-1}(U_{m:N})\Big)+\frac{m-k-1}{N}\Big\{\ \mu(\alpha,\beta)+$$
$$\frac{-\xi_\alpha+\mu(\alpha,\beta)}{\beta-\alpha}\left(\frac{k}{N+1}-\alpha\right)+\frac{\xi_\beta-\mu(\alpha,\beta)}{\beta-\alpha}\left(\frac{m}{N+1}-\beta\right)-$$
$$\frac{1}{\beta-\alpha}\left[\frac{1}{2f(\xi_\alpha)}-\frac{\mu(\alpha,\beta)-\xi_\alpha}{\beta-\alpha}\right]\frac{\frac{k}{N+1}(1-\frac{k}{N+1})}{N+2}+$$
$$\frac{1}{\beta-\alpha}\left[\frac{1}{2f(\xi_\beta)}+\frac{\mu(\alpha,\beta)-\xi_\beta}{\beta-\alpha}\right]\frac{\frac{m}{N+1}(1-\frac{m}{N+1})}{N+2}+$$
$$\left[\frac{\xi_\alpha+\xi_\beta-2\mu(\alpha,\beta)}{(\beta-\alpha)^2}\right]\frac{\frac{k}{N+1}(1-\frac{m}{N+1})}{N+2}+O(N^{-3/2})\Big\}-(\beta-\alpha)\mu(\alpha,\beta),$$

which easily leads to

$$\frac{1}{N}\Big\{\ \xi_\alpha(\alpha N-[\alpha N])-\xi_\beta(\beta N-[\beta N])-$$
$$\frac{1}{2f(\xi_\alpha)}\alpha(1-\alpha)+\frac{1}{2f(\xi_\beta)}\beta(1-\beta)\Big\}+O(N^{-3/2}). \qquad (A.3)$$

For B_1 we have

$$B_1=\frac{(\beta N-[\beta N])-(\alpha N-[\alpha N])}{[\beta N]-[\alpha N]}E\left(\frac{1}{N}\sum_{i=k}^{m}X'_{i:N}\right)=$$
$$\frac{(\beta N-[\beta N])-(\alpha N-[\alpha N])}{[\beta N]-[\alpha N]}\Big(\ (\beta-\alpha)\mu(\alpha,\beta)+\beta_N+O(N^{-3/2})\Big)$$
$$\frac{1}{N}\Big(\ (\beta N-[\beta N])-(\alpha N-[\alpha N])\Big)\mu(\alpha,\beta)+O(N^{-2}).$$

This together with (A.2)–(A.3) implies (A.1). The lemma is proved. □

Consider a trimmed mean T_N as in (3.1). Let $T_{N\Omega_k}$ is defined as in (1.8) of Putter and van Zwet (1998). We prove the following assertion.

LEMMA A.2. *Suppose that the conditions of Theorem 1.1 hold. Then*

$$\sum_{k=3}^{N}\binom{N-2}{k-2}ET^2_{N\Omega_k}=N^{-3}\left(\frac{\alpha^2(1-\alpha)^2}{f^2(\xi_\alpha)}+\frac{\beta^2(1-\beta)^2}{f^2(\xi_\beta)}\right)+o(N^{-3}) \qquad (A.4)$$

as $N\to\infty$.

Relation (A.4) directly yields that in the second condition of (1.18) in Theorem 1.2 of Putter and van Zwet (1998) is not satisfied for a Studentized trimmed mean, as Putter and van Zwet (1998) require that the l.h.s. of (A.4) is of order $N^{-7/2}$, instead of N^{-3} as in our relation (A.4).

PROOF. In Putter's Ph.D thesis (1994) it was proved that if T_N is a linear combination of order statistics, then

$$\sum_{k=3}^{N}\binom{N-2}{k-2}ET^2_{N\Omega_k}=E(Z_N-E(Z_N|U_{N-1},U_N))^2=$$

$$EZ_N^2 - E(T_{N,(1,2)})^2, \qquad (A.5)$$

(cf. (3.5.17), Putter (1994)), where $T_{N\Omega_k}, T_{N,(1,2)}$ are defined as in (1.8) of Putter and van Zwet (1998), Z_N is a r.v. defined as in (4.21) of van Zwet (1984), $U_1, \ldots, U_N$ are uniformly on (0,1) distributed r.v.'s. Let R_j denotes the rank of U_j among $U_1, \ldots, U_N$, $K_1 = R_{N-1} \wedge R_N$, $K_2 = R_{N-1} \vee R_N$. Take $X_{0:N} = -\infty$, $X_{N+1:N} = +\infty$ (cf. van Zwet (1984)). Let the functions G, H, M are defined as in (4.17) of van Zwet (1984), and define in addition the functions G_1 and H_1 by $G_1(x) = \int_{-\infty}^{x} F^2(y)\,dy$, $H_1(x) = \int_x^{\infty}(1-F(y))^2\,dy$. Then formula (4.21) of van Zwet (1984) reduces to

$$N^{1/2} Z_N = -\sum_{j=1}^{K_1}(c_{j+1}-c_j)(G_1(X_{j:N}) - G_1(X_{j-1:N}))+$$

$$\sum_{j=K_1}^{K_2-1}(c_{j+1}-c_j)(M(X_{j+1:N}) - M(X_{j:N})) - \sum_{j=K_2}^{N}(c_j - c_{j-1})(H_1(X_{j:N}) - H_1(X_{j+1:N})),$$

where in the trimmed mean case ($c_j = 1$ for $k \le j \le m$ and $c_j = 0$ for $j < k$, $j > m$) there are only two nonzero summands, which depend on K_1 and K_2. For instance, when $K_2 < k$ (which happens with probability $P(K_2 < k) = \alpha^2 + O(N^{-1})$), the value of $N^{1/2}Z_N$ equals

$$-[H_1(X_{k:N}) - H_1(X_{k+1:N})] + [H_1(X_{m+1:N}) - H_1(X_{m+2:N})] \stackrel{d}{=}$$

$$-[H_1 \circ F^{-1}(U_{k:N}) - H_1 \circ F^{-1}(U_{k+1:N})] + [H_1 \circ F^{-1}(U_{m+1:N}) - H_1 \circ F^{-1}(U_{m+2:N})],$$

where $U_{i:N}$ are order statistics of r.v.'s U_i, $i = 1, \ldots, N$. Application of a two term Taylor expansion of the function $H_1 \circ F^{-1}$ in neighborhoods of α and β respectively, together with the well-known facts that $E(s_i^2) = \frac{2}{(N+2)(N+1)}$, $E(s_i s_j) = \frac{1}{(N+2)(N+1)} (i \ne j)$, where $s_i = U_{i:N} - U_{i-1:N}$, $i = 1, \ldots, N$, yields that $E(Z_N^2 | K_2 < k) = \frac{2}{N^3}\left(\frac{(1-\alpha)^4}{f^2(\alpha)} + \frac{(1-\beta)^4}{f^2(\beta)} - \frac{(1-\alpha)^2(1-\beta)^2}{f(\alpha)f(\beta)}\right) + o(N^{-3})$, where $P(K_2 < k) = \alpha^2 + O(1/N)$. Analyzing in similar fashion the other possibilities for K_1 and K_2, we find that

$$EZ_N^2 = \frac{2}{N^3}\left(\frac{\alpha^2(1-\alpha)^2}{f^2(\xi_\alpha)} - \frac{\alpha^2(1-\beta)^2}{f(\xi_\alpha)f(\xi_\beta)} + \frac{\beta^2(1-\beta)^2}{f^2(\xi_\beta)}\right) + o(N^{-3}), \qquad (A.6)$$

as $N \to \infty$. Next we consider $T_{N,(1,2)}$. By formula (2.11) of Putter and van Zwet (1998) we have

$$N^{1/2} T_{N,(1,2)} = -\int_0^1 (I_{[U_1,1)}(t) - t)(I_{[U_2,1)}(t) - t)\binom{N-2}{k-2} t^{k-2}(1-t)^{N-k}\,dF^{-1}(t)+$$

$$\int_0^1 (I_{[U_1,1)}(t) - t)(I_{[U_2,1)}(t) - t)\binom{N-2}{m-1} t^{m-1}(1-t)^{N-m-1}\,dF^{-1}(t).$$

Define $\Delta F_{i,N}(x) = F_{i-1,N}(x) - F_{i,N}(x)$, where $F_{i,N}(x) = P(X_{i:N} \le x)$. Then the last relation implies that $E(T_{N,(1,2)})^2$ equals to

$$\frac{2}{N}\int_{-\infty}^{\infty}\int_{-\infty}^{z}\Big[-\int_{-\infty}^{\infty}(I_{[y,+\infty)}(x) - F(x))(I_{[z,+\infty)}(x) - F(x))\Delta F_{k-1,N-2}(x)\,dx+$$

$$\int_{-\infty}^{\infty}(I_{[y,+\infty)}(x) - F(x))(I_{[z,+\infty)}(x) - F(x))\Delta F_{m,N-2}(x)\,dx\Big]^2\,dF(y)\,dF(z) =$$

$$\frac{2}{N}\int_{-\infty}^{\infty}\int_{-\infty}^{z}[I_{k-1}(y,z)]^2\,dF(y)\,dF(z) + \frac{2}{N}\int_{-\infty}^{\infty}\int_{-\infty}^{z}[I_m(y,z)]^2\,dF(y)\,dF(z)-$$

$$\frac{4}{N}\int_{-\infty}^{\infty}\int_{-\infty}^{z}[I_{k-1}(y,z)I_m(y,z)]\,dF(y)\,dF(z), \tag{A.7}$$

where $I_r(y,z)$, $r=k-1,m$, is defined as

$$\int_{-\infty}^{y}\Delta F_{r,N-2}(x)\,dG_1(x)-\int_{y}^{z}\Delta F_{r,N-2}(x)\,dM(x)-\int_{z}^{\infty}\Delta F_{r,N-2}(x)\,dH_1(x).$$

Consider the first term at the r.h.s. of (A.7) (the treatment of the second and third term is similar). Integrating by parts, we reduce it to

$$\frac{2}{N}\int_{-\infty}^{\infty}\int_{-\infty}^{z}\Big[(G_1(y)+M(y))\Delta F_{k-1,N-2}(y)+(H_1(z)-M(z))\Delta F_{k-1,N-2}(z)\ -$$

$$-\int_{-\infty}^{y}G_1(x)\,d(\Delta F_{k-1,N-2}(x))+\int_{y}^{z}M(x)d(\Delta F_{k-1,N-2}(x))+$$

$$\int_{z}^{\infty}H_1(x)d(\Delta F_{k-1,N-2}(x))\Big]^2\,dF(y)\,dF(z).$$

Note that the 'basic' support of the function $\Delta F_{k-1,N-2}(x)=F_{k-2,N-2}(x)-F_{k-1,N-2}(x)$ is some interval $I_\alpha(A)=[\xi_\alpha-A(\log N/N)^{1/2},\xi_\alpha+A(\log N/N)^{1/2}]$ in the sense that for every $c>2$ we have the following bound: $\sup_{y\in R\setminus I_\alpha(A)}\Delta F_{k-1,N-2}(y)=O\left(P\left(|U_{k:N}-\alpha|>(\log N/N)^{1/2}\right)\right)=O(N^{-c})$, where $A>0$ is some constant, depending only on c, α and $f(\xi_\alpha)$. Moreover, smoothness conditions imply that $\sup_{y\in I_\alpha(A)}\Delta F_{k-1,N-2}(y)=O(N^{-1})$ as $N\to\infty$. Thus, the last expression reduces to

$$\frac{2}{N}\int_{-\infty}^{\infty}\int_{-\infty}^{z}\Big[-\int_{-\infty}^{y}G_1(x)\,d(\Delta F_{k-1,N-2}(x))+\int_{y}^{z}M(x)d(\Delta F_{k-1,N-2}(x))+$$

$$\int_{z}^{\infty}H_1(x)d(\Delta F_{k-1,N-2}(x))\Big]^2\,dF(y)\,dF(z)+o(N^{-3}), \tag{A.8}$$

as $N\to\infty$. Consider the integrand in (A.8) and note that if $I_\alpha(A)\subset(-\infty,y)$, then integrand equals to $[E(G_1(X_{k-2:N-2})-G_1(X_{k-1:N-2}))]^2+o(N^{-2})=\frac{\alpha^4}{N^2}\frac{1}{f^2(\xi_\alpha)}+o(N^{-2})$, and the corresponding part of the integral in (A.8) (in the domain where $Y=\min(X_1,X_2)\ge\xi_\alpha$) equals to $\frac{(1-\alpha)^2\alpha^4}{f^2(\xi_\alpha)}N^{-3}+o(N^{-3})$. Argueing similarly for the cases $I_\alpha(A)\subset(y,z)$ and $I_\alpha(A)\subset(z,+\infty)$ (the cases $y\in I_\alpha(A)$ or $z\in I_\alpha(A)$ are negligible) we obtain that the quantity (A.8), and hence the first term at the r.h.s. in (A.7), equals to $\left(\frac{\alpha^4(1-\alpha)^2}{f^2(\xi_\alpha)}+2\frac{\alpha^2(1-\alpha)^2}{f^2(\xi_\alpha)}+\frac{(1-\alpha)^4\alpha^2}{f^2(\xi_\alpha)}\right)N^{-3}+o(N^{-3})=\frac{\alpha^2(1-\alpha)^2}{f^2(\xi_\alpha)}N^{-3}+o(N^{-3})$. Similarly for the second term at the r.h.s. of (A.7) we get $\frac{\beta^2(1-\beta)^2}{f^2(\xi_\beta)}N^{-3}+o(N^{-3})$, and for the third one we obtain $-2\frac{\alpha^2(1-\beta)^2}{f(\xi_\alpha)f(\xi_\beta)}N^{-3}+o(N^{-3})$. Together these results give us

$$E(T_{N,(1,2)})^2=N^{-3}\left(\frac{\alpha^2(1-\alpha)^2}{f^2(\xi_\alpha)}-2\frac{\alpha^2(1-\beta)^2}{f(\xi_\alpha)f(\xi_\beta)}+\frac{\beta^2(1-\beta)^2}{f^2(\xi_\beta)}\right)+o(N^{-3}) \tag{A.9}$$

as $N\to\infty$. The relations (A.5), (A.6) and (A.9) imply (A.4) and the lemma is proved. □

Acknowledgement

This paper was completed while the first author visited CWI in January-February, 2002; we thank the Netherlands Organisation for Scientific Research (NWO) for making this visit to CWI possible.

References

[1] BAHADUR, R.R. (1966). A note on quantiles in large samples, Ann. Math. Statist. **37** 577 - 580.

[2] BENTKUS, V., GÖTZE F. and VAN ZWET W.R. (1997) An Edgeworth expansion for symmetric statistics. Ann. Statist. **25** 851 - 896.

[3] BICKEL, P.J., GÖTZE F. and VAN ZWET W.R. (1986) The Edgeworth expansion for $U-$ statistics of degree 2. Ann. Statist. **14** 1463 - 1484.

[4] BJERVE, S. (1974). Error bounds and asymptotic expansions for linear combinations of order statistics. Unpublished Ph.D. thesis, Berkeley.

[5] HALL, P. and PADMANABHAN, A.R. (1992). On the bootstrap and the trimmed mean. Journal of Multivariate Analysis **41**, 132 - 153.

[6] HELMERS, R. (1979). Edgeworth expansions for trimmed linear combinations of order statistics. In Proceedings, Second Prague Symp. Asymptotic Statist. (P.Mandl and M.Hušková, Eds), pp. 221 - 232. North-Holland, Amsterdam

[7] HELMERS, R. (1991). On the Edgeworth expansion and the bootstrap approximation for a Studentized $U-$statistic. Ann. Statist. **19** 470 - 484.

[8] HOEFFDING, W. (1963) Probabilities inequalities for sum of bounded random variables. J. Amer.Statist.Assoc. **58**, no. 301, 13 - 30.

[9] PUTTER, H. (1994). Consistency of resampling methods. Ph.D. thesis, Leiden University.

[10] PUTTER, H. and VAN ZWET, W.R. (1998). Empirical Edgeworth expansions for symmetric statistics, Ann. Statist. **26** 1540 - 1569.

[11] REISS, R.-D. (1989). Approximate Distributions of Order Statistics *With Applications to Nonparametric Statistics*. Springer-Verlag, New York.

[12] VAN ZWET, W.R. (1984). A Berry – Esseen bound for symmetric statistics, Z. Wahrsch. Verw. Gebiete **66** 425 - 440.