Centrum Wiskunde & Informatica

**REPORT**RAPPORT

PNA

Probability, Networks and Algorithms

***Probability, Networks and Algorithms***

PNA

# On the *M* fewer than *N* bootstrap approximation to the trimmed mean

N.V. Gribkova, R. Helmers

**REPORT PNA-E0810 DECEMBER 2008**

Centrum Wiskunde & Informatica (CWI) is the national research institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organisation 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

# On the *M* fewer than *N* bootstrap on the *M* fewer than *N* bootstrap approximation to the trimmed mean

**Abstract**

*We show that the $M$ fewer than $N$ ($N$ is the real data sample size, $M$ denotes the size of the bootstrap resample; $M/N \to 0$, as $M \to \infty$) bootstrap approximation to the distribution of the trimmed mean is consistent without any conditions on the population distribution $F$, whereas Efron's naive (i.e. $M = N$) bootstrap as well as the normal approximation fails to be consistent if the population distribution $F$ has gaps at the two quantiles where the trimming occurs.*

*2000 Mathematics Subject Classification:* 62E20;62G30;62G09;60F05
*Keywords and Phrases:* *M* fewer than *N* bootstrap approximation, consistency, asymptotic normality; trimmed mean

# On the $M$ Fewer Than $N$ Bootstrap Approximation to the Trimmed Mean

N.V. Gribkova[1] and R. Helmers[2]

[1] *St.Petersburg State University, Mathematics and Mechanics Faculty, 198504, St.Petersburg, Stary Peterhof, Universitetsky pr. 28, Russia*
*E-mail: nvg@rol.ru*
[2] *Centre for Mathematics and Computer Science*
*P.O.Box 94079, 1090 GB Amsterdam, The Netherlands*
*E-mail: helmers@cwi.nl*

**Abstract**

*We show that the $M$ fewer than $N$ ($N$ is the real data sample size, $M$ denotes the size of the bootstrap resample; $M/N \to 0$, as $M \to \infty$) bootstrap approximation to the distribution of the trimmed mean is consistent without any conditions on the population distribution $F$, whereas Efron's naive (i.e. $M = N$) bootstrap as well as the normal approximation fails to be consistent if the population distribution $F$ has gaps at the two quantiles where the trimming occurs.*

2000 Mathematics Subject Classification: Primary 62E20, 62G30, 62G09; Secondary 60F05.

*Key words and phrases*: $M$ fewer than $N$ bootstrap approximation, consistency, asymptotic normality, trimmed mean.

## 1 Introduction

The trimmed mean is a well known estimator of a location parameter, the main reason for applying it in statistics is robustness. The centre of a distribution is often estimated by the sample mean or the sample median. However, it is well known that the sample mean is sensitive to outliers and thus not robust. On the other hand, the sample median is robust against outliers but not very efficient if the underlying distribution is, for instance, normal. An estimator showing intermediate behavior, and which includes both the sample mean and sample median, is the trimmed (sample) mean (cf. [9], [11]). Compared with robust $M$-estimates of maximum likelihood type, the trimmed mean not only has the same asymptotic variance but also is easy to compute.

The limit distribution of the trimmed mean for an arbitrary population distribution was found by Stigler [12]. Specifically he has shown that in order for the trimmed mean to be asymptotically normal, it is necessary and sufficient that the sample is trimmed at sample quantiles for which the corresponding population quantiles are uniquely defined. His result was extended to the case of slightly trimmed mean (when the fraction of trimmed data is vanishing when $N$ gets large) by Csörgő et.al [5].

In two recent papers by Gribkova and Helmers [6]–[7] the validity of the one-term Edgeworth expansion for a (Studentized) trimmed mean and bootstrapped trimmed mean were established and simple explicit formulas of the first leading terms of these expansions were obtained. Using these expansions the second order correctness of the $M$ out of $N$ bootstrap approximation ($N$ being the real data sample size, $M$ denotes the size of the bootstrap resample) was obtained, provided the trimmed mean is properly defined and a local smoothness condition near the quantiles where the trimming occurs is satisfied. We also assume a relation between $M$ and $N$, as $\min(M, N)$ gets large.

In this article we prove that, without any smoothness condition on the underlying population distribution function $F$, the distribution of the bootstrapped trimmed mean tends to the same limit as the distribution of the trimmed mean, whenever both the $M$ out of $N$ bootstrapped trimmed mean and the trimmed mean are properly normalized and $M/N \to 0$, as $M \to \infty$. This implies that the $M$ fewer than $N$ bootstrap approximation is always consistent for the trimmed mean when data are i.i.d., including non smooth cases, with 'gaps' in the underlying distribution, i.e. intervals of positive length near the quantiles where the trimming occurs with $F$-measure zero, when the normal approximation and the Efron's naive (i.e. $M = N$) bootstrap approximation do not work.

Some general results concerning the $M$ fewer than $N$ bootstrap with replacement and without replacement can be found in [2]. In the present article we establish that the $M$ fewer than $N$ bootstrap with replacement works for the case of trimmed means, also when the ordinary Efron bootstrap fails. Our main result – Theorem 1 – supplements the very general Theorem 2 in Bickel, Gotze and van Zwet [2]. Our proof of Theorem 1 is relatively simple: it is based on approximating the trimmed mean and its $M$ fewer than $N$ bootstrap counterpart by sums of i.i.d. Winsorized random variables.

Theorem 1 presents a new example, useful in statistical practice, where the $M$ fewer than $N$ bootstrap with replacement works, while Efron bootstrap fails. We discuss an application of Theorem 1 by constructing a $M \ll N$-bootstrap confidence interval for the parameter $\mu(\alpha, \beta)$ – the population trimmed mean (cf.(1.3)) – in section 3. A numerical example is given in Section 5.

We refer to [2] for some other realistic examples of $M$ fewer than $N$ bootstrap success and ordinary bootstrap failure. A possible alternative way to prove assertion ($ii$) of our Theorem 1 would consist of checking the general assumptions of Theorem 2 of [2] for the special case of the trimmed mean.

Consider a sequence $X_1, X_2, \dots$ of independent and identically distributed (i.i.d.) real-valued random variables (r.v.) with common distribution function (*df*) $F$, and let $X_{1:N} \le \dots \le X_{N:N}$ ($N = 1, 2, \dots$) be the order statistics corresponding to a sample $X_1, \dots, X_N$ of size $N$ from $F$.

Let $F^{-1}(u) = \inf\{x : F(x) \ge u\}$, $0 < u \le 1$, denotes the left-continuous inverse function of the *df* $F$ and put $F_N^{-1}$ to be the inverse of $F_N$, the empirical distribution of $(X_1, \dots, X_N)$.

Define two versions of the trimmed mean by

$$T_{N_{(1)}} = \frac{1}{([\beta N] - [\alpha N])} \sum_{i=[\alpha N]+1}^{[\beta N]} X_{i:N}\,, \qquad T_{N_{(2)}} = \frac{1}{\beta - \alpha} \int_\alpha^\beta F_N^{-1}(u)\, du, \tag{1.1}$$

where $0 < \alpha < \beta < 1$ are any fixed numbers, $[\cdot]$ represents the greatest integer function. Note that $T_{N_{(1)}}$ is the natural definition of a trimmed mean; and $T_{N_{(2)}}$ (the 'plug-in version' of a trimmed mean) — only slightly different from the classical $T_{N_{(1)}}$ — is very useful in a bootstrap context, in particular, it enables us to obtain the second order correctness property for the $M$ out of $N$ bootstrap for a (Studentized) trimmed mean $T_{N_{(2)}}$ (cf. [7]).

Let $X_1^*, \dots, X_M^*$ be a bootstrap resample of the size $M = M(N)$ from the empirical *df* $F_N$ based on the first $N$ original observations $X_1, \dots, X_N$; $F_M^*$ denotes the bootstrap empirical *df*, and $X_{1:M}^* \le \dots \le X_{M:M}^*$ — the corresponding order statistics. Here and throughout this article we use the shorthand notation $M$, omitting its argument $N$.

Define the $M$ out of $N$ bootstrap counterparts of the trimmed means by:

$$T^*_{N,M_{(1)}} = \frac{1}{([\beta M]-[\alpha M])} \sum_{i=[\alpha M]+1}^{[\beta M]} X^*_{i:M}\,, \quad T^*_{N,M_{(2)}} = \frac{1}{\beta-\alpha}\int_\alpha^\beta (F_M^*)^{-1}(u)\,du, \quad (1.2)$$

where $\alpha$ and $\beta$ are same as in (1.1).

It is well known that both versions of the trimmed mean can be used as statistical estimators of the location parameter

$$\mu(\alpha,\beta) = \frac{1}{\beta-\alpha}\int_\alpha^\beta F^{-1}(u)\,du, \quad (1.3)$$

the population trimmed mean. The bootstrap counterpart of the parameter $\mu(\alpha,\beta)$ is given by $\mu_N(\alpha,\beta) = \frac{1}{\beta-\alpha}\int_\alpha^\beta F_N^{-1}(u)\,du = T_{N_{(2)}}$.

Let us introduce the $\nu$-th $(0<\nu<1)$ quantile of of $F$, $F_N$ and $(F_M^*)$: $\xi_\nu = F^{-1}(\nu)$, $\xi_{\nu N:N} = F_N^{-1}(\nu)$, $\xi^*_{\nu M:M} = (F_M^*)^{-1}(\nu)$, and note that the following simple relations are valid:

$$\begin{aligned} T_{N_{(2)}} &= c_{\alpha,N}\left(T_{N_{(1)}} - \xi_{\alpha N:N}\right) + T_{N_{(1)}} - c_{\beta,N}\left(T_{N_{(1)}} - \xi_{\beta N:N}\right), \\ T^*_{N,M_{(2)}} &= c_{\alpha,M}\left(T^*_{N,M_{(1)}} - \xi^*_{\alpha M:M}\right) + T^*_{N,M_{(1)}} - c_{\beta,M}\left(T^*_{N,M_{(1)}} - \xi^*_{\beta M:M}\right), \end{aligned} \quad (1.4)$$

with $c_{\nu,\kappa} = \frac{\nu\kappa-[\nu\kappa]}{(\beta-\alpha)\kappa} = O(1/\kappa)$, where $\kappa = N, M$ and $\nu = \alpha, \beta$. Since $\xi_{\nu N:N}$ and $T_{N_{(j)}}$, $j=1,2$, are bounded uniformly in $N$ with probability one (a.s.) as $N\to\infty$ and in view of (1.4)

$$N^{1/2}(T_{N_{(2)}} - T_{N_{(1)}}) \to 0, \quad \text{a.s.,} \quad \text{as } N\to\infty. \quad (1.5)$$

Moreover, since $\xi^*_{\nu M:M}$, $\nu=\alpha,\beta$, and $T^*_{N,M_{(j)}}$, $j=1,2$ are bounded uniformly with $P^*$-probability of the order $1-O(M^{-c})$, for every $c>0$, a.s. [P], (cf. [7]),

$$M^{1/2}(T^*_{N,M_{(2)}} - T^*_{N,M_{(1)}}) \to_{P^*} 0, \quad a.s., \quad \text{as} \quad \min(N,M)\to\infty. \quad (1.6)$$

Here and throughout this paper $P^*$ denotes bootstrap probability measure having discrete mass points $X_{i:N}$ with atoms $1/N$, and $E^*$ denotes the corresponding expectation.

The relations (1.5) – (1.6) imply that the limit distributions of both versions of the trimmed mean, given by (1.1), are identical; the same remark is valid for the corresponding bootstrap counterparts. Therefore, we will use in the sequel with impunity the following shorthand notation:

$$T_N = T_{N_{(j)}}, \qquad T^*_{N,M} = T^*_{N,M_{(j)}}\,, \qquad j = 1,\,2.$$

Our results on the limit distributions and the $M$ fewer than $N$ bootstrap correctness will be proved for the trimmed mean given as $T_{N_{(2)}}$ and its bootstrap counterpart

$T^*_{N,M_{(2)}}$; but in view of the previous remarks these results are also valid for $T_{N_{(1)}}$, and its bootstrap counterpart $T^*_{N,M_{(1)}}$.

We conclude this section by noting that the second order asymptotic properties (one-term Edgeworth expansions) of two versions of the trimmed mean and the corresponding bootstrap counterparts are different, and that $T_{N_{(2)}}$ is the more appropriate definition of a trimmed mean in a bootstrap context (cf. Gribkova and Helmers [7]).

# 2 The asymptotic distribution of $T_N$ and $T^*_{N,M}$ and the $M \ll N$ bootstrap consistency

Following Stigler [12] we define two quantities

$$A = \sup\{x: \ F(x) \le \alpha\} - F^{-1}(\alpha)\ ; \quad B = \sup\{x: \ F(x) \le \beta\} - F^{-1}(\beta)\,,$$

which are both equal to zero if and only if the inverse function $F^{-1}$ is continuous at the two points $\alpha$ and $\beta$.

Define auxiliary r.v.'s $Y_i$, having distribution function $G(x)$:

$$Y_i = \begin{cases} X_i + A, & X_i \le \xi_\alpha, \\ X_i, & \xi_\alpha < X_i \le \xi_\beta, \\ X_i - B, & \xi_\beta < X_i. \end{cases} \qquad G(x) = \begin{cases} F(x-A), & x < \xi_\alpha + A, \\ F(x), & \xi_\alpha + A \le x < \xi_\beta, \\ F(x+B), & \xi_\beta \le x, \end{cases} \tag{2.1}$$

$i = 1, \dots, N$. Define the inverse function $G^{-1}(u) = \inf\{x: \ G(x) \ge u\}$ and note that this function is continuous at the points $\alpha$ and $\beta$. Define $\alpha$ and $\beta$-th quantiles of $G$:

$$\eta_\alpha = G^{-1}(\alpha) = \xi_\alpha + A\,; \quad \eta_\beta = G^{-1}(\beta) = \xi_\beta\,. \tag{2.2}$$

Let us introduce r.v.'s $Y_i$ Winsorized outside $(\eta_\alpha, \eta_\beta]$, in other words

$$W_i = \eta_\alpha \vee (Y_i \wedge \eta_\beta)\,, \quad i = 1, \dots, N\,, \tag{2.3}$$

where $a \wedge b = \min(a,b)$ and $a \vee b = \max(a,b)$, the corresponding quantile function is

$$Q(u) = \eta_\alpha \vee (G^{-1}(u) \wedge \eta_\beta)\,.$$

Define

$$\mu_W = \int_0^1 Q(u)\,du, \quad \sigma_W^2 = \int_0^1 (Q(u) - \mu_W)^2\,du, \tag{2.4}$$

the first two central moments of $W_i$. Note that $\mu_(\alpha,\beta) = \frac{1}{\beta-\alpha}\int_\alpha^\beta G^{-1}(u)\,du$. (cf. (1.3))

Let $(Z_1, Z_2, Z_3)$ be a normally distributed $\mathcal{N}(\mathbb{O}, \mathbb{C})$ random vector with zero expectation and the covariance matrix

$$\mathbb{C} = \begin{pmatrix} \alpha(1-\alpha) & \alpha(\eta_\alpha - \mu_W) & -\alpha(1-\beta) \\ \alpha(\eta_\alpha - \mu_W) & \sigma_W^2 & (1-\beta)(\eta_\beta - \mu_W) \\ -\alpha(1-\beta) & (1-\beta)(\eta_\beta - \mu_W) & \beta(1-\beta) \end{pmatrix} \tag{2.5}$$

Define a r.v.

$$Z = (\beta - \alpha)^{-1}\Big( - A\max(0, Z_1) + Z_2 + B\max(0, Z_3)\Big)\,, \tag{2.6}$$

with expectation $\mathbb{E}Z = (\sqrt{2\pi}(\beta - \alpha))^{-1}\big( - A(\alpha(1-\alpha))^{1/2} + B(\beta(1-\beta))^{1/2}\big)$ (cf. Stigler [12]). Using (2.6) and the covariance matrix (2.5) we easily obtain the second moment of the r.v. $Z$: $\mathbb{E}Z^2 = (\beta-\alpha)^{-2}\Big\{A^2\frac{\alpha(1-\alpha)}{2} + \sigma_W^2 + B^2\frac{\beta(1-\beta)}{2}$ $-A\alpha(\eta_\alpha - \mu_W) + B(1-\beta)(\eta_\beta - \mu_W) - AB\frac{\alpha(1-\beta)}{\pi}\Big(\sqrt{\frac{\beta-\alpha}{\alpha(1-\beta)}} - arctg\Big(\sqrt{\frac{\beta-\alpha}{\alpha(1-\beta)}}\Big)\Big)\Big\}$.

Let $F_Z$ denotes the *df* of the r.v. $Z$. We will suppose that the following condition of the non degeneracy is satisfied: $\eta_\alpha \neq \eta_\beta$ (i.e., $\xi_\alpha + A$ is not an atom of the *df* $F$ with probability mass at least $(\beta-\alpha)$; note that this equivalent to assuming $\sigma_W > 0$).

The following Theorem is our main result, which consist of two parts: first of all a version of the classical theorem by Stigler [12] on the asymptotic distribution of the trimmed mean, and secondly the $M$ fewer than $N$ bootstrap analogue of that theorem.

**Theorem 1**. *Suppose that $\sigma_W > 0$. Then*

$$(i) \quad \sup_{x\in\mathbb{R}}\Big|P\big(N^{1/2}(T_N - \mu(\alpha,\beta)) \le x\big) - F_Z(x)\Big| \longrightarrow 0\,, \ \ as \ \ N \to \infty \ ;$$

$$(ii) \quad \sup_{x\in\mathbb{R}}\Big|P^*\big(M^{1/2}(T^*_{N,M} - T_N) \le x\big) - F_Z(x)\Big| \longrightarrow 0\,, \ \ as \ \ M \to \infty\ ,\ M/N \to 0\ ,$$

*in $P$-probability. Moreover,*

$(iii)$ *if* $(M \log\log N)/N \to 0$, *as* $M \to \infty$, *then the relation* $(ii)$ *is valid* $a.s.$ $[P]$.

The proof of Theorem 1 is relegated to the Section 4. Note that the first relation $(i)$ is nothing but the result by Stigler [12]. Nevertheless, we will give a short proof of Stigler's Theorem as our proof is slightly different from one given in [12]; it is based on the stochastic approximation which was also used in [6]-[7] — a sum of i.i.d. Winsorized r.v.'s.. Note also that the representation of the limit r.v. $Z$ (but not the limit distribution itself) differs from one given in [12].

It is clear from our proof of the relations $(ii)$–$(iii)$ in Theorem 1 (cf. Section 3) that the condition $M/N \to 0$ is necessary for the convergence to the limit distribution only when the gaps $A$ or/and $B$ are not equal to zero. However, when $A = B = 0$, the limit distribution is the normal one under the sole condition that $N \wedge M \to \infty$.

We should also note that the condition $\sigma_W > 0$ imposed in Theorem 1 is not needed for the convergence in $(i)$–$(iii)$ (cf. proof of Theorem 1). However, if $\eta_\alpha = \eta_\beta$, then $F_Z$ can be discontinuous and the convergence in $(i)$–$(iii)$ is not uniform anymore (cf. [12]). However, this exceptional case does not seem to be interesting in statistical applications.

The following corollary is a direct consequence of Theorem 1.

**Corollary 1**. *Suppose that $\sigma_W > 0$, then*

$$(i) \quad \sup_{x\in\mathbb{R}}\Big|\, P^*\big(M^{1/2}(T^*_{N,M} - T_N) \le x\big) - P\big(N^{1/2}(T_N - \mu(\alpha,\beta)) \le x\big)\Big| \longrightarrow 0,$$

*in $P-$probability, as* $M \to \infty$, $M/N \to 0$ .

$(ii)$ *Relation* $(i)$ *is valid a.s.* $[P]$, *as* $M \to \infty$, *if* $(M \log\log N)/N \to 0$.

Inference on trimmed means using the normal approximation will be invalid when gaps are present near the quantiles where the trimming occur, for instance, when one has to deal with grouped data (cf. Stigler [12]). The $M$ fewer than $N$ bootstrap approximation, however, also works when normality fails, provided only that at $\eta_\alpha$ there is not an atom with mass at least $\beta - \alpha$ of the distribution $F$.

Theorem 1 and its corollary cover a case where the classical bootstrap is not consistent , but the $M$ fewer than $N$ bootstrap is consistent, without assuming any (smoothness) condition on the underlying distribution of the observations. We note that in a way Theorem 1 belongs to the third case of the three cases distinguished in Bickel and Sakov [3].

The question remains how well the $M$ fewer than $N$ bootstrap approximation approximates the true distribution in finite samples; the rate of convergence might be quite slow and will depend on the choice of $M$. The authors intend to pursue this matter elsewhere.

# 3 Some applications

Since the limit distribution of $N^{1/2}(T_N - \mu(\alpha,\beta))$ has a non zero expectation $\mathbb{E}Z$ in general case, it is natural to estimate the parameter $\mu(\alpha,\beta)$ by $T_N - N^{-1/2}\widehat{\mathbb{E}Z}$ (cf. (3.1) below). The asymptotic bias $\mathbb{E}Z$ depends on the unknown quantities $A$ and $B$ (cf. (2.6)) and we need a consistent estimate of $\mathbb{E}Z$. Fortunately, we can apply the $M$ fewer than $N$ bootstrap procedure for this purpose. Define

$$\widehat{\mathbb{E}Z} = M^{1/2}\mathbb{E}^*(T^*_{N,M} - T_N), \qquad \widehat{\sigma^2}_{N,M} = M\left(\mathbb{E}^*(T^*_{N,M} - T_N)^2\right) - \widehat{\mathbb{E}Z}^2. \tag{3.1}$$

The following result means the consistency of these $M$ fewer than $N$ bootstrap estimators of the expectation and the variance of the limit distribution $F_Z$.

**Proposition 1**. *Suppose that $\sigma_W > 0$, then*

$$\widehat{\mathbb{E}Z} - \mathbb{E}Z \longrightarrow 0\,, \qquad\qquad \widehat{\sigma^2}_{N,M}/\mathbb{V}ar(Z) \longrightarrow 1\,, \tag{3.2}$$

*in $P-$probability, as $M \to \infty,\ \ M/N \to 0$. Moreover, relations* (3.2) *are valid a.s.* $[P]$, *as $M \to \infty,\ \ (M \log\log N)/N \to 0$.*

The proof of the Proposition 1 is relegated to the Section 4. Note that both $\mathbb{E}^*(T^*_{N,M})$ and $\widehat{\sigma^2}_{N,M}$ can be easily computed using Monte Carlo.

We now apply our results to construct a confidence interval for the parameter $\mu(\alpha,\beta)$. We suppose that $\xi_\alpha + A \neq \xi_\beta$, that is the natural (in the context of estimating) condition of absence of degeneracy is satisfied. Take any $p > 0$ close to zero. Then we obtain the following asymptotic $M$ fewer than $N$ bootstrap confidence interval for the population trimmed mean:

$$T_N \ - \frac{\hat{t}_{(1-\frac{p}{2})} + \widehat{\mathbb{E}Z}}{\sqrt{N}} \ \leq\ \mu(\alpha,\beta)\ \leq\ T_N - \frac{\hat{t}_{\frac{p}{2}} + \widehat{\mathbb{E}Z}}{\sqrt{N}}, \tag{3.3}$$

with the coverage probability close to $1-p$, where $\hat{t}_\gamma$ $(\gamma = \frac{p}{2},\ (1-\frac{p}{2}))$ are the $\gamma$-th quantiles of the *df* of $M^{1/2}(T^*_{N,M} - \mathbb{E}^*(T^*_{N,M}))$, which can be computed by Monte Carlo method.

As a consequence of Theorem 1 and Proposition 1 we obtain the following result for a Studentized trimmed mean:

**Corollary 2**. *Suppose that $\sigma_W > 0$, then*

$$\sup_{x\in\mathbb{R}} \left| P\left(N^{1/2}(T_N - \mu(\alpha,\beta)) \leq \widehat{\sigma}_{N,M}\ x\right) - F_Z(\widehat{\sigma}_{N,M}\ x)\right| \longrightarrow 0, \tag{3.4}$$

*in P-probability, as $M \to \infty$, $M/N \to 0$.*

To obtain the corresponding confidence interval based on the Studentized statistic, we need the $M$ fewer than $N$ bootstrap counterpart of (3.4). This can be established with the aid of the double bootstrap procedure, where we estimate $\widehat{\sigma}_{N,M}$ using the bootstrap resamples of size $m << M$ from the bootstrap empirical *df* $F^*_M$. However, to obtain the double bootstrap approximation one will need extensive Monte Carlo simulations.

# 4 Proofs

In this section we prove Theorem 1 and Proposition 1 stated in section 2.

**Proof of Theorem 1**. First we prove the relation $(i)$. Rewrite $T_N$ in terms of $Y_{i:N}$ – the order statistics corresponding to the sample $Y_1, \ldots, Y_N$, with the empirical *df* $G_N$. Define the binomial r.v.'s $N_\nu = \sharp\{i : X_i \le \xi_\nu\} = \sharp\{i : Y_i \le \eta_\nu\}$, $\nu = \alpha, \beta$. Let $\eta_{\nu N:N} = G_N^{-1}(\nu)$ denotes the $\nu$-th sample quantile, $\nu = \alpha, \beta$. Then we can write

$$(\beta-\alpha)T_N = -A\,\frac{N_\alpha - \alpha N}{N}\mathbb{I}_{\{N_\alpha \ge \alpha N\}} + (\beta-\alpha)T_N(Y) + B\,\frac{\beta N - N_\beta}{N}\mathbb{I}_{\{N_\beta \le \beta N\}}, \quad (4.1)$$

where $\mathbb{I}_E$ denotes the indicator of the event $E$, and $(\beta-\alpha)T_N(Y) = -\frac{\alpha N - [\alpha N]}{N}\,\eta_{\alpha N:N} + \frac{1}{N}\sum_{i=[\alpha N]+1}^{[\beta N]} Y_{i:N} + \frac{\beta N - [\beta N]}{N}\,\eta_{\beta N:N}$. Consider an average of i.i.d. r.v.'s $W_i$: $\overline{W}_N = \frac{1}{N}\sum_{i=1}^{N} W_i = \frac{N_\alpha}{N}\,\eta_\alpha + \frac{1}{N}\sum_{i=N_\alpha+1}^{N_\beta} Y_{i:N} + \frac{N-N_\beta}{N}\,\eta_\beta$, with $\mathbb{E}\overline{W}_N = \mu_W$ (cf. (2.4)).

We have

$$\begin{aligned}&(\beta-\alpha)(T_N(Y) - \mu(\alpha,\beta)) - (\overline{W}_N - \mu_W)\\ &= -\,\frac{\alpha N - [\alpha N]}{N}\,(\eta_{\alpha N:N} - \eta_\alpha) + \frac{\beta N - [\beta N]}{N}\,(\eta_{\beta N:N} - \eta_\beta) + S_{\alpha,N} - S_{\beta,N},\end{aligned} \quad (4.2)$$

where with $\nu = \alpha, \beta$ and $S_{\nu,N} = \frac{sign(N_\nu - [\nu N])}{N}\sum_{i=([\nu N]\wedge N_\nu)+1}^{N_\nu \vee [\nu N]}(Y_{i:N} - \eta_\nu)$, $sign(0) = 0$. Obviously $N^{1/2}\,|S_{\nu,N}| \le \frac{|N_\nu - [\nu N]|}{N^{1/2}}\,\big(\big|\,\eta_{\nu N:N} - \eta_\nu)\big| \vee \big|\,Y_{N_\nu:N} - \eta_\nu)\big|\big)$.

Since $\frac{|N_\nu - [\nu N]|}{N^{1/2}}$ is bounded in probability and $\eta_{\nu N:N} - \eta_\nu \to 0$, $Y_{N_\nu:N} - \eta_\nu \to 0$ (because the inverse $G^{-1}$ of the *df* of $Y_i$ is continuous at the points $\alpha$ and $\beta$), we obtain the following relation

$$N^{1/2}(\beta-\alpha)(T_N(Y) - \mu(\alpha,\beta)) = N^{1/2}(\overline{W}_N - \mu_W) + o_P(1) \quad (4.3)$$

Define random variables $Z_{1,N} = \frac{N_\alpha - \alpha N}{N^{1/2}}$, $Z_{2,N} = N^{1/2}(\overline{W}_N - \mu_W)$, $Z_{3,N} = \frac{\beta N - N_\beta}{N^{1/2}}$.

Then relations (4.1), (4.3) together imply

$$N^{1/2}(\beta-\alpha)(T_N - \mu(\alpha,\beta)) = -A\,(\,0 \vee Z_{1,N}) + Z_{2,N} + B\,(\,0 \vee Z_{3,N}) + o_P(1).$$

The random vector $(Z_{1,N}, Z_{2,N}, Z_{3,N})$ has zero expectation and covariance matrix $\mathbb{C}$ for each $N$, and the Central Limit Theorem for sums of i.i.d. r.v.'s implies $(i)$.

We will prove the relations $(ii)$ and $(iii)$ simultaneously, and our proof is based on the same arguments adapted to the bootstrap world. Let us introduce the auxiliary r.v.'s

$$Y_i^* = \begin{cases} X_i^* + A, & X_i^* \le \xi_\alpha,\\ X_i^*, & \xi_\alpha < X_i^* \le \xi_\beta,\\ X_i^* - B, & \xi_\beta < X_i^*.\end{cases} \quad (4.4)$$

$i = 1, \dots, M$, which represents a bootstrap resample from $Y_1, \dots, Y_N$. Let $Y^*_{i:M}$ denote the corresponding order statistics, $\eta^*_{\nu M:M}$ – the $\nu$-th bootstrapped quantile, $\nu = \alpha,\, \beta$.

Denote $M^*_\nu = \sharp\{i : X^*_i \le \xi_\alpha,\, i = 1, \dots, M\}$, $\nu = \alpha,\, \beta$. Then we can rewrite $T^*_{N,M}$ in terms of $Y^*_{i:M}$

$$(\beta-\alpha)T^*_{N,M} = -A\,\frac{M^*_\alpha - \alpha M}{M}\mathbb{I}_{\{M^*_\alpha \ge \alpha M\}} + (\beta-\alpha)T^*_{N,M}(Y) + B\,\frac{\beta M - M^*_\beta}{M}\mathbb{I}_{\{M^*_\beta \le \beta M\}}, \tag{4.5}$$

where $(\beta-\alpha)T^*_{N,M}(Y) = -\frac{\alpha M - [\alpha M]}{M}\,\eta^*_{\alpha M:M} + \frac{1}{M}\sum_{i=[\alpha M]+1}^{[\beta M]} Y^*_{i:M} + \frac{\beta M - [\beta M]}{M}\,\eta^*_{\beta M:M}$. First we write

$$M^{1/2}(\beta-\alpha)(T^*_{N,M}(Y) - T_N) = M^{1/2}(\beta-\alpha)(T^*_{N,M}(Y) - T_N(Y)) + R_{N,M}, \tag{4.6}$$

where

$$\begin{aligned} R_{N,M} &= M^{1/2}(\beta-\alpha)(T_N(Y) - T_N) \\ &= M^{1/2}A\,\max(\,0, \frac{N_\alpha - \alpha N}{N}) - M^{1/2}B\,\max(\,0, \frac{\beta N - N_\beta}{N}) \end{aligned} \tag{4.7}$$

(cf. (4.1)). Note that $\frac{\nu N - N_\nu}{N^{1/2}}$ $(\nu = \alpha\,,\beta)$ is bounded in probability, and as $(M/N)^{1/2}$ tends to zero, $R_{N,M}$ tends to zero in probability, this fact will be use to prove $(ii)$. Moreover, note that $\frac{\nu N - N_\nu}{(\log\log\, N\, N)^{1/2}}$ is bounded $a.s.$ and if $((M \log\log\, N)/N)^{1/2}$ tends to zero, $R_{N,M}$ tends to zero $a.s.$ $[P]$, what we will used to prove $(iii)$.

Consider $Y^*_i$ Winsorized outside of $(\eta_{\alpha N:N}, \eta_{\beta N:N}]$: $W^*_i = \eta_{\alpha N:N} \vee (Y^*_i \wedge \eta_{\beta N:N})$, $i = 1, \dots, M$, and put $N^*_\nu = \sharp\{i : Y^*_i \le \eta_{\nu N:N}, i = 1, \dots, M\}$, $\nu = \alpha\,\beta$. Define the average $\overline{W}^*_{N,M}$, its expectation is equal to $\mathbb{E}^*\overline{W}^*_{N,M} = \frac{[\alpha N]}{N}\eta_{\alpha N:N} + \frac{1}{N}\sum_{i=[\alpha N]+1}^{[\beta N]} Y_{i:N} + \frac{N-[\beta N]}{N}\eta_{\beta N:N}$. Then for the first term at the r.h.s. of (4.6) we have

$$\begin{aligned} &(\beta-\alpha)(T^*_{N,M}(Y) - T_N(Y)) - (\overline{W}^*_{N,M} - \mathbb{E}^*\overline{W}^*) = \\ &-\frac{\alpha M - [\alpha M]}{M}(\eta^*_{\alpha M:M} - \eta_{\alpha N:N}) + \frac{\beta M - [\beta M]}{M}(\eta^*_{\beta M:M} - \eta_{\beta N:N}) + (S^*_{\alpha,M} - S^*_{\beta,M}) \end{aligned} \tag{4.8}$$

where for $\nu = \alpha\,\beta$ $S^*_{\nu,M} = \frac{sign(N^*_\nu - [\nu M])}{M}\sum_{i=([\nu M]\wedge N^*_\nu)+1}^{N^*_\nu \vee [\nu M]}(Y^*_{i:M} - \eta_{\nu N:N})$, and we have the following estimate

$$M^{1/2}\left|S^*_{\nu,M}\right| \le \frac{|N^*_\nu - [\nu M]|}{M^{1/2}}\left(\left|\,\eta^*_{\nu M:M} - \eta_{\nu N:N})\right| \vee \left|\,Y^*_{N^*_\nu:M} - \eta_{\nu N:N})\right|\right). \tag{4.9}$$

Since $\frac{|N^*_\nu - [\nu M]|}{M^{1/2}}$ is bounded in $P^*-$probability $a.s.$ $[P]$ and $\eta^*_{\nu M:M} - \eta_{\nu N:N} \to 0$, $Y^*_{N^*_\nu:M} - \eta_{\nu N:N} \to 0$ (because the inverse function $G^{-1}$ is continuous at the points $\alpha$ and $\beta$ (cf. Gribkova & Helmers (2007))), we obtain

$$\begin{aligned} &M^{1/2}(\beta-\alpha)(T^*_{N,M} - T_N) \\ =&M^{1/2}(\overline{W}^*_{N,M} - \mathbb{E}^*\overline{W}^*_{N,M}) - A\,\max\Big(0, \frac{M^*_\alpha - \alpha M}{M^{1/2}}\Big) + B\,\max\Big(0, \frac{\beta M - M^*_\beta}{M^{1/2}}\Big) \\ &+ R_{N,M} + R^*_{N,M}, \end{aligned} \tag{4.10}$$

where $R^*_{N,M}$ tends to zero in $P^*-$probability $a.s.$ $[P]$ (cf. (4.8), (4.9)) and $R_{N,M}$ as in (4.7).

Define r.v.'s $Z^*_{1,M} = \frac{N^*_\alpha - \alpha M}{M^{1/2}}$, $Z^*_{2,M} = M^{1/2}(\overline{W}^*_{N,M} - \mathbb{E}^*\overline{W}^*_{N,M})$, $Z^*_{3,M} = \frac{\beta M - N^*_\beta}{M^{1/2}}$. Then relation (4.10) implies

$$M^{1/2}(\beta-\alpha)(T^*_{N,M} - T_N) = -A \max\Big(0, \big(Z^*_{1,M} + \frac{M^*_\alpha - N^*_\alpha}{M^{1/2}}\big)\Big) + Z^*_{2,M}$$

$$+B \max\Big(0, \big(Z^*_{3,N} + \frac{N^*_\beta - M^*_\beta}{M^{1/2}}\big)\Big) + R_{N,M} + R^*_{N,M}. \tag{4.11}$$

The random vector $(Z^*_{1,M}, Z^*_{2,M}, Z^*_{3,M})$ tends in the bootstrap distribution to $\mathcal{N}(\mathbb{O}, \mathbb{C})$, where $\mathbb{C}$ is the covariance matrix given by (2.5), when $M \to \infty$ (cf. Bickel & Freedman [1]), the remainder term $R_{N,M} + R^*_{N,M}$ tends to zero in $P^*$-probability on a set with $P$-probability $1-\epsilon$, for any $\epsilon > 0$, when $M/N \to 0$, as $M \to \infty$, and it tends to zero $a.s.$ $[P]$, when $(M \log\log N)/N \to 0$ , as $M \to \infty$ (cf. arguments below (4.7)).

It remains to note that conditionally on $X_1, \ldots, X_N$ the r.v.'s $|M^*_\nu - N^*_\nu|$, $\nu = \alpha, \beta$, have the binomial distribution with parameters $\left(\frac{|N_\nu - [\nu N]|}{N}, M\right)$ respectively, and an application of Chebyshev inequality directly yields $\frac{|M^*_\nu - N^*_\nu|}{M^{1/2}} = \frac{|N_\nu - [\nu N]|}{N} M^{1/2} + o_{P^*}(1) = \frac{|N_\nu - [\nu N]|}{N^{1/2}} \left(\frac{M}{N}\right)^{1/2} + o_{P^*}(1)$, whereas the latter quantity tends to zero in $P^*$-probability on a set with $P$-probability $1-\epsilon$, for any $\epsilon > 0$, when $M/N \to 0$, as $M \to \infty$ (which implies $(ii)$), and $a.s.$ $[P]$ when $(M \log\log N)/N \to 0$, as $M \to \infty$ (which implies $(iii)$). The theorem is proved.

**Proof of Proposition 1**. We begin our proof with relation (4.10), which directly yields

$$M^{1/2}(\beta-\alpha)\mathbb{E}^*\left(T^*_{N,M} - T_N\right) = -A\,\mathbb{E}^*\Big(\max\big(0, \frac{M^*_\alpha - \alpha M}{M^{1/2}}\big)\Big)$$

$$+B\,\mathbb{E}^*\Big(\max\big(0, \frac{\beta M - M^*_\beta}{M^{1/2}}\big)\Big) + R_{N,M} + \mathbb{E}^*\left(R^*_{N,M}\right), \tag{4.12}$$

where $R_{N,M} \to 0$ in $P-$probability when $M \to \infty$, $M/N \to 0$, whereas this convergence is a.s. $[P]$, if $M \log\log N/N \to 0$ (cf. proof of Theorem 1).

Next we note that $R^*_{N,M} \to 0$ in $P^*-$probability a.s. $[P]$ when $N \wedge M \to \infty$. Since the r.v. $R^*_{N,M}$ is uniformly integrable (cf. [4]) because it has a finite variance (in $P^*-$measure) a.s. $[P]$, we can conclude that $\mathbb{E}^*\big(R^*_{N,M}\big) \to 0$ a.s. $[P]$ as $N \wedge M \to \infty$. Finally, we note that the r.v.'s $\frac{M^*_\alpha - \alpha M}{M^{1/2}}$ and $\frac{\beta M - M^*_\beta}{M^{1/2}}$ tend in the bootstrap distribution to the r.v.'s $Z_1$ and $Z_3$ respectively, in $P-$probability when $M \to \infty$, $M/N \to 0$, and a.s. $[P]$, if $M \log\log N/N \to 0$. It remains to note that the function $f(x) = \max(0, x)$ is continuous and since r.v.'s $\frac{M^*_\alpha - \alpha M}{M^{1/2}}$ and $\frac{\beta M - M^*_\beta}{M^{1/2}}$ have the finite variances a.s. $[P]$, they are uniformly integrable. Thus, we can conclude that (cf. (4.12)) $M^{1/2}\mathbb{E}^*\big(T^*_{N,M} - T_N\big) \longrightarrow \mathbb{E}Z$, in $P-$probability when $M \to \infty$, $M/N \to 0$, and a.s. $[P]$, if $M \log\log N/N \to 0$, $M \to \infty$.

The proof of the second relation in (3.2) is similar. The proposition is proved.

# 5 A numerical example

In this section we supplement our results by some simulations. We consider an example, where the underlying distribution is the uniform one on the set $S = [-9, -5] \cup [-4, 4] \cup$

[23, 27] with the density $f(x) = 1/16$, $x \in S$, and $f(x) = 0$, $x \in \mathbb{R} \setminus S$.

Take $\alpha = 1 - \beta = 0.25$, then the population trimmed mean $\mu(\alpha, \beta) = 0$ (cf. (1.3)), $\xi_\alpha = -5$, $\eta_\alpha = -4$, $\xi_\beta = \eta_\beta = 4$, $A = 1$, $B = 19$. We produce the sample of size $N = 1000$ of the underlying r.v. $K = 10^6$ times, and compute the value of $N^{-1/2}(T_N - \mu(\alpha, \beta))$ in each case. As result we obtain a Monte Carlo sample of size $K$ for this r.v. Then given a sample of size $N$ from the initial *df*, we perform the bootstrap simulations and obtain the Monte Carlo samples of size $10^5$ for the r.v. $M^{-1/2}(T^*_{N,M} - T_N)$, with $M = N = 1000$, $M = N/5 = 200$ and $M = N/20 = 50$.

Finally, we obtain a Monte Carlo sample from the distribution of r.v. $Z$ using $K$ independent realizations of the normal distributed random vector $(Z_1, Z_2, Z_3)$ with given covariance matrix $\mathbb{C}$ (cf. (2.5)), and applying formula (2.6) to obtain a value of $Z$.

In Figure 1 the graphs of the empirical densities are presented. We plot a histogram (step-line) corresponding the Monte Carlo sample of $N^{-1/2}(T_N - \mu(\alpha, \beta))$. Then we plot graph of density, corresponding to the normal r.v. $\mathbb{N}(a, \sigma^2)$, where $a$ and $\sigma^2$ are computed as the empirical average and the empirical variance derived from the Monte Carlo sample of $N^{-1/2}(T_N - \mu(\alpha, \beta))$. Finally, we plot three histograms (smoothed by splines), corresponding to the Monte Carlo samples of the limit r.v. $Z$ and bootstrap r.v.'s $M^{-1/2}(T^*_{N,M} - T_N)$, with $M = N = 1000$ and $M = N/20 = 50$.

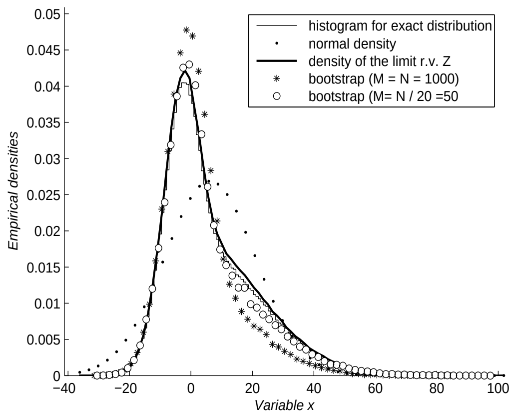


Figure 1: Behavior of the densities.

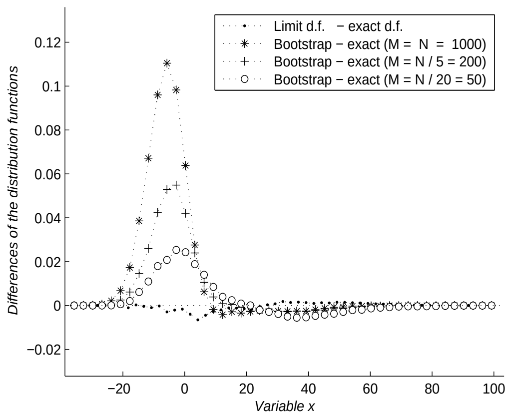


Figure 2: Deviations of the *df*'s

In Figure 2 we plot a graph of the deviation: *df* $F_Z$ minus exact *df* of the r.v. $N^{-1/2}(T_N - \mu(\alpha, \beta)))$, where $F_Z$ (the *df* of the limit r.v. $Z$) as well as the exact *df* are produced on the basis of the corresponding Monte Carlo samples. Moreover, we plot the graphs of the deviations: *df* of the r.v. $M^{-1/2}(T^*_{N,M} - T_N)$ minus exact *df*, where $M_1 = N = 1000$, $M_2 = N/5 = 200$, $M_3 = N/20 = 50$.

The simulations confirm and illustrate our result that the normal approximation as well as the naive (when $M = N$) bootstrap are not valid for the trimmed mean in general situation when the inverse $F^{-1}$ of the population *df* has the jumps at the points $\alpha$ and/or $\beta$, and we can see also that the $M$ fewer than $N$ bootstrap rectifies the failure of the classical Efron's bootstrap.

Take $1 - p = 0.95$ be our nominal confidence level. Then we obtain the following results on the confidence interval for the parameter $\mu(\alpha, \beta) = 0$ in our simulations: $T_N = 0.0101$, the 95-percentage confidence interval based on the exact *df* – $[-1.2717,\ 0.4736]$ , obtained by Monte Carlo using $10^6$ samples of size 1000. And with $M = N = 1000$, $M = N/5 = 200$, $M = N/20 = 50$ respectively, we obtain estimates $T_N - N^{-1/2}\widehat{\mathbb{E}Z}$ equal to $-0.0307$, $-0.1008$ and $-0.1444$, the corresponding $M$ fewer than $N$ bootstrap 95-percentage confidence intervals given by (3.3)) are : $[-0.8148, 0.4487]$, $[-1.1343,\ 0.4592]$ and $[-1.3368,\ 0.4510]$, their estimated coverage probabilities are 0.8590, 0.9290, 0.9490, respectively.

We remark that if $F^{-1}$ is continuous at the points $\alpha$ and $\beta$ (that is $A = B$ – the $iff$ condition for asymptotic normality of the trimmed mean [12]), then $G \equiv F$, $Y_i = X_i$, $i = 1, \ldots, N$, (cf. (2.1)), and $Y_j^* = X_j^*$, $j = 1, \ldots, M$ (cf. (4.4)), the limit distribution in Theorem 1 is the normal one with asymptotic variance $(\beta - \alpha)^2\sigma_W^2$ (cf. (2.4), (2.5)); in this case the plug-in estimate of the asymptotic variance (cf. [6], [7]) is consistent as well as its $M$ out of $N$ ($M$, $N$ vary independently, $N \wedge M \to \infty$) bootstrap version. So, the $M$ out of $N$ bootstrap works like normality works under the Stigler's $iff$ condition, both for trimmed mean and Studentized trimmed mean, as $N \wedge M \to \infty$.

To conclude this paper we note also that a completely different way of approximating *df* of the trimmed mean accurately is to use saddlepoint approximations. These approximations will presumably work better in the far tail of the distribution in comparison with the $M$ fewer than $N$ bootstrap approximation considered in the present paper. Recently these saddlepoint approximations were established for the trimmed mean and the Studentized trimmed mean under suitable conditions. We refer to [10] for more details.

# References


[1] *Bickel P.J. and Freedman D.A.* Some asymptotic theory for the bootstrap. — Ann. Statist., 1981, v. 9, p. 1196-1217.

[2] *Bickel P.J., Götze F. and van Zwet W.R.* Resampling fewer than $n$ observations: gains, losses and remedies for losses. — Statistica Sinica, 1997, v. 7, p. 1-31.

[3] *Bickel P.J. and Sakov A.* Extrapolation and the bootstrap. — Sankhya, 2002, v. 64, Ser. A, Pt. 3 , p. 640-652.

[4] *Billingsley P. Convergence of Probability Measures*, Wiley, New York, 1968.

[5] *Csörgő S., Haeusler E. and Mason D.M.* The asymptotic distribution of trimmed sums. — Ann. Probab., 1988, v. 16, p. 672-699.

[6] *Gribkova N.V. and Helmers R.* The empirical Edgeworth expansion for a Studentized trimmed mean. — Math. Meth. Statist., 2006, v. 15, p. 61-87.

[7] *Gribkova N.V. and Helmers R.* On the Edgeworth expansion and the $M$ out of $N$ bootstrap accuracy for a Studentized trimmed mean. — Math. Meth. Statist., 2007, v. 16, p. 142-176.

[8] *Hall P. and Padmanabhan A.R.* On the bootstrap and the trimmed mean.— J. of Multivariate Analysis, 1992, v. 41, p. 132-153.

[9] *Hampel F.R., Ronchetti E.M., Rousseeuw P.J. and Stahel W.A. Robust Statistics. The Approach Based on Influence Functions*, John Wiley & Sons, New York, 1986.

[10] *Helmers R., Jing B.-Y., Qin G. and Zhou W.* Saddlepoint approximations to the trimmed mean. — Bernoulli, 2004, v. 10, p. 465-501.

[11] *Huber P.J. Robust Statistics*, John Wiley & Sons, New York, 1981.

[12] *Stigler S.M.* The asymptotic distribution of the trimmed mean. — Ann. Statist., 1973, v. 1, p. 472-477.