Asymptotic Mean Squared Errors for Asymmetrically Trimmed Means

Gary W. Oehlert

University of Minnesota
School of Statistics

Technical Report #445

December 1984

University of Minnesota
School of Statistics
Department of Applied Statistics
St. Paul, Minnesota 55108

Asymptotic Mean Squared Errors for
Asymmetrically Trimmed Means

Gary W. Oehlert
Dept. of Applied Statistics
University of Minnesota

ABSTRACT

An asymmetric trimmed mean with trimming only on the right can be a consistent estimate of the mean if the trimming fraction goes to zero. We show under mild regularity conditions that the mean squared errors of such trimmed means are asymptotically larger that the mean squared error of the sample mean.

AMS 1980 subject classification: 62G05, 62G20

Key words: Asymmetric trimmed mean, asymptotic expansion, mean square error

Asymptotic Mean Squared Errors for Asymmetrically Trimmed Means

Gary W. Oehlert

1. Introduction and Summary. Let $X_1 < X_2 < \ldots < X_n$ be the ordered observations of an independent, identically distributed sample from a distribution F, where F is absolutely continuous with finite mean μ and variance σ^2, $F(0) = 0$ and $F^{-1}(1) = \infty$. We wish to estimate μ. The distribution F is unknown, so that parametric procedures are unavailable. The sample mean $\bar{X}$ is the natural estimate of μ, but the sample mean is not robust. A simple robust alternative to $\bar{X}$ is an asymmetrically trimmed mean $\bar{X}_k$, where

$$\bar{X}_k = \frac{1}{n-k} \sum_{i=1}^{n-k} X_i \quad .$$

If k/n (with k allowed to depend on n) goes to zero, then $\bar{X}_k$ is a consistent estimate of μ. We employ asymmetric trimming because we believe that any lack of robustness is due to large, not small, values. Symmetric trimmed means provide good variance reduction over the sample mean for long tailed distributions in the usual robustness setup (see, for example, Bickel 1965), so it might be hoped that asymmetric trimmed means would provide good mean square error (MSE) improvement over the sample mean in asymmetric situations. The purpose of this note is to show that this MSE reduction does not occur. Under fairly mild restrictions, the bias introduced through trimming is greater than the variance reduction gained, so that the MSE of the trimmed mean is greater than σ^2/n, even for long tailed distributions like the Pareto.

Our approach will be asymptotic, and our goal is to prove the following:

Theorem. Let $X_1 < X_2 < \ldots < X_n$ be an ordered sample from F, where F is absolutely continuous with finite variance, $F(0) = 0$, and $F^{-1}(1) = \infty$.

If 1) there exist $\alpha > 2$ and X_0 such that $X^{\alpha}(1-F(x))$ is decreasing for $X > X_0$, and 2) $k\ (=k(n)) = o(n)$ and $\log(n) = O(k)$,

then $E(\bar{X}_k - \mu)^2 = \frac{\sigma^2}{n} + \left(\int_{q(k,n)}^{\infty} x\, dF\right)^2$ + smaller orders ,

where $q(k,n) = \inf\{X\colon F(X) \geq \frac{n-k}{n}\}$. If F also satisfies:

$h = F^{-1}$ is twice continuously differentiable with

$\limsup_{t \uparrow 1} (1-t)\ |h''(t)/h'(t)| < 5/4$

and $\limsup_{t \uparrow 1} (1-t)|h'(t)/h(t)|$ bounded, then the growth

restriction $\log(n) = O(k)$ can be relaxed to $k \to \infty$.

The two sets of assumptions are used to show that the tail order statistics behave like their quantiles, in particular, to show that

$$E\left| \frac{X_{n-k}}{q(k,n)} - 1 \right|^4 \to 0 .$$

The first set of conditions makes stronger assumptions about k and weaker assumptions about the tail of F than the second set. The first assumptions are still met if $X^{\alpha}f(X)\sim C$ for $3 < \alpha < 5$, while the second set requires $\alpha \geq 5$. The price we pay for working with the longer tails under the first set of assumptions is that we must trim off more tail values for the asymptotic MSE expansion to be valid.

A procedure related to trimming is truncation. Define the trucated mean to be

$$\bar{y}_k = \frac{1}{n} \sum_{i=1}^{n} X_i I(X_i \leq q(k,n)) ,$$

where $I(\cdot)$ is the indicator function. What we in fact show in proof of the theorem is that under our assumptions, the MSE of the trimmed mean is equal to the MSE of the truncated mean to the given order of approximation.

To illustrate the theorem, consider the Pareto family $F(X) = 1 - 1/X^{\alpha-1}$ for $X \geq 1$. This is a long tailed distribution, and we might have expected trimmed means to produce substantial MSE reduction. However, the conclusion of the theorem is that the MSE of a trimmed mean for the Pareto should be

$$\frac{\frac{\alpha-1}{\alpha-3} - \left(\frac{\alpha-1}{\alpha-2}\right)^2}{n} + \frac{\alpha-1}{\alpha-2}^{\,2} \left(\frac{k}{n}\right)^{2\,\frac{\alpha-2}{\alpha-1}} + \text{smaller orders.}$$

For $\alpha = 4$, this is

$$\frac{\sigma^2}{n}\left(1 + 3\,\frac{k^{4/3}}{n^{1/3}} + \text{smaller orders}\right) .$$

If we take $k = n^{1/10}$, then our two term expansion predicts a doubling of the MSE for n up to 243.

2. Proof of the Theorem. We begin with two lemmas which will establish the quantile approximation to the tail order statistics. For the rest of this paper, we will denote $q(k,n)$ simply by q.

Lemma 1. Suppose $X_1 < X_2 < \ldots < X_n$ is an ordered sample from F, where F is absolutely continuous, $F^{-1}(1) = \infty$, and there exist $\alpha > 0$ and X_0 such that $X^\alpha(1-F(x))$ is decreasing for $X > X_0$. If $k = o(n)$ and $\log(n) = O(k)$, then

$$E \left| \frac{X_{n-k}}{q(k,n)} - 1\right|^P \to 0 \qquad \text{for positive } P .$$

Proof. It suffices to show that $X_{n-k}/q \to 1$ in probability and that $|X_{n-k}/q|^P$ is uniformly integrable. Let $X_i = F^{-1}(u_i)$, where u_i is the i-th order statistic from a sample of n uniform (0,1) random variables.

By using the fact that $X^\alpha(1-F(x))$ is decreasing in the tails, it is not difficult to show that

$$F^{-1}(u) \le [(1-u)(1-F(x))^{-1}]^{-1/\alpha} x ,$$

for $F^{-1}(u) \ge x > x_0$. From this, we get for $u_{n-k} \ge \frac{n-k}{n}$

$$[1 + |u_{n-k} - \frac{n-k}{n}| \frac{n}{k}]^{-1/\alpha} \le \frac{X_{n-k}}{q} \le [1 - |u_{n-k} - \frac{n-k}{n}|\frac{n}{k}]^{-1/\alpha} .$$

The same can be shown for $u_{n-k} \le \frac{n-k}{n}$. Since

$$|u_{n-k} - \frac{n-k}{n}|\frac{n}{k} = 0_p(\frac{1}{\sqrt{k}}) ,$$

we conclude that $X_{n-k}/q \to 1$ in probability.

To show uniform integrability of $|X_{n-k}/q|$, we must show that for given $\varepsilon > 0$, there exists a B such that for $b > B$,

$$q^{-1} \int_{bq}^{\infty} |y/q|^P \, dG_{n-k,n} (y) < \varepsilon ,$$

where $G_{n-k,n}$ is the distribution of X_{n-k}. By standard order statistics results and the assumption that $y^P(1-F(y))^k$ is decreasing for some $k_0 > 0$, we need to bound

$$c_1 q^{-P-1} \frac{n!}{(n-k-1)!\ k!} \int_{bq}^{\infty} F(y)^{n-k-1} (1-F(y))^{k-k_0} dF(y)$$

$$\le c_1 q^{-P-1} \frac{n!}{(n-k_0)!} \frac{(k-k_0)\ !}{k!} (1-I_{F(bq)}(n-k, k-k_0 + 1)) ,$$

where $I_{\cdot}(\cdot,\cdot)$ is the incomplete beta function. Continuing, the integral of interest is bounded by

$$c_1 q^{-P-1} \frac{n!}{(n-k_0)!} \frac{(k-k_0)!}{k!} I_{1-F(bq)}(k-k_0 + 1, n-k)$$

$$\le c_1 q^{-P-1} \frac{n!}{(n-k_0)!} \frac{(k-k_0)!}{k!} I_{\frac{k}{n} b^{-\alpha}} (k-k_0 + 1, n-k) \text{ for } b>1 .$$

Using the standard expansion of the incomplete beta as a sum, one can show that for b large enough the last incomplete beta is bounded by

$$C_2 \frac{(n-k_0)!}{(k-k_0+1)! \quad (n-k-1)!} \left(\frac{k}{n} b^{-\alpha}\right)^{k-k_0+1} \left(1-\frac{k}{n} b^{-\alpha}\right)^{n-k-1}$$

Substituting this back in and using Stirling's approximation, our integral of interest is now bounded by

$$C_3 \, q^{-P-1} \frac{1}{\sqrt{k}} \left(\frac{n}{k}\right)^{k_0} b^{-\alpha k} \, e^{k(1-b^{-\alpha})} \quad .$$

By choosing b large enough, this can be made as small as desired, so that uniform integrability and the lemma are proved.

A second lemma which may be used to prove convergence of $X_{n-k/q}$ follows.

<u>Lemma 2</u>. Let $0 < p \leq r$ be real numbers, and let $h = F^{-1}$ be twice continuously differentiable on (0,1) and satisfy

$$\limsup_{t \to 1} (1-t)|h''(t)/h'(t)| < 1 + 1/r \quad \text{and}$$

$$\limsup_{t \to 1} (1-t)|h'(t)/h(t)| \quad \text{bounded}$$

If $k = o(n)$ and $k \to \infty$, then

$$E \left|\frac{X_{n-k}}{q} - 1\right|^P \to 0 \quad .$$

<u>Proof</u>. Lemma 2 is a direct application of Lemma A2.3 of Albers, Bickel and Van Zwet (1976).

Proof of Theorem: We use the following shorthand notation:

$$A = \bar{X}_k - \frac{n-k}{n} \bar{X}_k \quad ,$$

$$B = \frac{n-k}{n} \bar{X}_k - \frac{1}{n} \sum_{i=1}^{n} X_i' \quad ,$$

and $$C = \frac{1}{n} \sum_{i=1}^{n} X_i' - \mu \quad ,$$

where $X_i' = X_i I(X_i \leq q)$ are truncated X_i's. The MSE of $\bar{X}_k$ is then the expected value of $(A + B + C)^2$.

Clearly, $E(A^2) = 0\left(\frac{k^2}{n^2}\right)$. Next,

$$|B| = \frac{1}{n} \left| \sum_{j \varepsilon J} X_j \right| \quad ,$$

where J is the index set of the order statistics in the difference of the two sums in B. For $j \varepsilon J$,

$$|x_j - q| \leq |x_{n-k} - q|, \text{ so that}$$

$$|B| \leq \frac{|J|}{n} (q + |X_{n-k} - q|) \quad ,$$

where $|J|$ is the cardinality of J. By Cauchy-Schwartz, either of the lemmas, and the central moments of a binomial, we see that

$$E(B^2) = 0\left(\frac{k}{n^2} q^2\right) \quad .$$

The cross product term E(AB) is $0\left(\frac{k^{3/2}}{n^2} q\right)$ by Cauchy-Schwartz. The $E(C^2)$ term is just the MSE of $\bar{y}_k$.

$$E(C^2) = \frac{1}{n}\left[\int_0^q x^2 dF - \left(\int_0^q x dF\right)^2\right] + \left(\int_q^\infty x\, dF\right)^2$$

$$= \frac{\sigma^2}{n} + \left(\int_q^\infty x\, dF\right)^2 + \text{smaller orders.}$$

The A term is equal to $\frac{k}{n-k}(C + \mu + B)$, so that

$$E(AC) = O\left(\frac{k}{n^2}\right) + o\left(E(BC)\right) + O\left(\frac{k}{n}\int_q^\infty x\, dF\right) .$$

The BC term may be written

$$BC = q/n\left[n-k-\sum_{i=1}^{n} I(X_i \le q)\right](\bar{y}_k - \mu)$$

$$+ \frac{1}{n}\sum_{j\varepsilon J}|X_j - q|\,(\bar{y}_k - \mu) \quad .$$

By conditioning on $\sum_{i=1}^{n} I(X_i \le q)$, the first term in BC can be shown to have expectation $O(\frac{k\,q}{n^2})$.

Let $L = n-k+1$ if the index set J is empty and $L = \min_{j\varepsilon J} j$ otherwise.

With L defined thusly, $\sum_{i=1}^{L-2} x_i$ and $\sum_{j\varepsilon J}|x_j - q|$ are conditionally independent given X_{L-1} . The second term in E(BC) is

$$\frac{1}{n}\left|E\left\{\sum_{j\varepsilon J}|X_j - q|\left[\frac{1}{n}\sum_{i=1}^{L-1} x_i - \mu + \frac{1}{n}\sum_{i=L}^{L+|J|-1} x_i\, I(x_{n-k} \le q)\right]\right\}\right|$$

$$\le \frac{1}{n}\left|E\left\{\sum_{j\varepsilon J}|x_j - q|\left(\frac{1}{n}\sum_{i=1}^{L-1} x_i - \mu\right)\right\}\right| + o\left(\frac{1}{n^2}q^2 k\right) ,$$

by Cauchy-Schwartz and the Lemmas. Continuing, condition on L, $|J|$, and X_{L-1} and use conditional independence to get

$$\frac{1}{n} E[\ |J|\ |x_{n-k} - q| \qquad |E(X|X < X_{L-1}) - \mu| + \frac{g}{n} + \mu|\frac{L-2-n}{n}|\ |] + o(\frac{1}{n^2} q^2 k)$$

$$\leq o(\frac{k^{1/2}}{n} q)\ \{E^{1/2}(E(X|X < X_{L-1}) - \mu)^2 + \frac{g}{n} + O(\frac{k}{n})\} + o(\frac{1}{n^2} q^2 k)\ .$$

To finish the proof, we must evaluate $E(E(X|X < X_{L-1}) - \mu)^2$. For large z, we have

$$|E(X|X \leq z) - \mu| \leq C \int_z^\infty x\, dF\ ,$$

which by our assumptions on $1-F(x)$ can be bounded by $C_2\, z(1-F(z))$.

Since $X_{L-1} \leq q$, we have

$$E(E(X|X < X_{L-1}) - \mu)^2 \leq q^2 E(1-u_{L-1})^2\ ,$$

where $u_{L-1} = F(X_{L-1})$ is the L-1st order statistic for the n uniforms corresponding to the X's. Note that

$$1-u_{L-1} \leq 1 - \frac{n-k}{n} + |u_{n-k} - \frac{n-k}{n}|\ ,\ \text{so that}$$

$$E(1-u_{L-1})^2 = O(\frac{k^2}{n^2})\ .$$

Recombining all BC terms, we see that $E(BC) = o(\frac{k^{3/2}}{n^2} q^2)$.

Checking the orders of all terms, we see that

$$E(\bar{X}_n - \mu)^2 = \frac{\sigma^2}{n} + (\int_q^\infty x\, dF)^2 + \text{smaller orders},$$

and the theorem is proved.

REFERENCES

Albers, W., Bickel, P., and Van Zwet, W. R. (1976). **Asymptotic expansions for the power of distribution free tests in the one-sample problem.** Ann. Statist. 4, 108-156.

Bickel, P. J. (1965). **On some robust estimates of location.** Ann. Math. Statist. 36, 847-858.

David, H. A. (1981). Order Statistics. Wiley, New York.