The Annals of Statistics
1986, Vol. 14, No. 1, 348–354

MAXIMUM ASYMPTOTIC VARIANCES OF TRIMMED MEANS UNDER ASYMMETRIC CONTAMINATION[1]

BY JOHN R. COLLINS

University of Calgary

We consider the following problem arising in robust estimation theory: Find the maximum asymptotic variance of a trimmed mean used to estimate an unknown location parameter when the error distribution is subject to asymmetric contamination. The model for the error distribution is $F = (1-\varepsilon)F_0 + \varepsilon G$, where F_0 is a known distribution symmetric about 0, ε is fixed proportion of contamination, and G is an unknown and possibly asymmetric distribution. We prove, under the assumption that F_0 has a symmetric unimodal density function f_0, that the maximal asymptotic variance is obtained when G places mass 1 at either $+\infty$ or $-\infty$. The key idea of the proof is first to maximize the asymptotic variance subject to the side conditions $F(a) = \alpha$ and $F(b) = 1 - \alpha$ when a and b are given.

1. Introduction and summary. Let $X_1, \ldots, X_n$ be a random sample from a distribution $F(x - \theta)$, where θ is an unknown parameter to be estimated. Let $T_\alpha = T_\alpha[X_1, \ldots, X_n]$ denote the α-trimmed mean as defined, e.g., on page 58 of Huber (1981). Then under mild regularity conditions on F, $n^{1/2}[T_\alpha - ET_\alpha]$ converges in distribution to a normal distribution with mean 0 and variance $V(F)$, where (ref. Andrews et al. (1972), pages 31 and 34):

$$(1)\quad V(F) = \frac{1}{(1-2\alpha)^2}\left\{\int_a^b (x - c(\alpha))^2\, dF + \alpha\left[(a - c(\alpha))^2 + (b - c(\alpha))^2\right]\right\},$$

where

$$(2)\qquad c(\alpha) = \int_a^b x\, dF + \alpha(a + b),$$

and where $a = F^{-1}(\alpha)$ and $b = F^{-1}(1-\alpha)$.

A problem arising in robust estimation theory is to evaluate the supremum of $V(F)$ as F varies over distributions of the form

$$(3)\qquad F = (1-\varepsilon)F_0 + \varepsilon G,$$

where F_0 is a fixed known distribution symmetric about 0, G is unknown, and ε is a fixed proportion of contamination. Here the constants ε and α are required to satisfy $0 < \varepsilon < \alpha < \frac{1}{2}$ in order to avoid breakdown. It is well known that when the unknown contaminating distribution G is restricted to be symmetric about 0, then T_α is an unbiased estimator of θ and $V(F)$ is maximized by the symmetric distribution G which places mass $\frac{1}{2}$ at each of $+\infty$ and $-\infty$.

Received February 1985; revised June 1985.
[1]This research was supported by the Natural Sciences and Engineering Research Council of Canada under Grant A-4499.
AMS 1980 *subject classifications*. Primary 62F35; secondary 62F12.
Key words and phrases. Trimmed means, asymmetry, asymptotic variances, robust estimation.

Consider now the situation when the unknown contaminating distribution G in (3) is *not* required to be symmetric. Although T_α is not an unbiased estimator of θ under asymmetric contamination, the problem of maximizing $V(F)$ remains of interest. (See Section 4.9 of Huber (1981) for further motivation for this problem.) Huber (1981) considered the case where F_0 is a normal distribution and made the "highly plausible" conjecture that $V(F)$ is maximized over all F of form (3) when G places mass 1 at either $+\infty$ or $-\infty$.

In this paper we prove that Huber's conjecture is true whenever the fixed F_0 in (3) has a density f_0 which is symmetric about 0 and unimodal. The main difficulty in proving the result is that the limits of integration in formula (1) depend on F. The device used to circumvent this difficulty is to first maximize $V(F)$ over all F of form (3) subject to the side conditions $F(a) = \alpha$ and $F(b) = 1 - \alpha$. A simple argument using the method of moment spaces yields the maximum asymptotic variance, $V(a, b)$, subject to $F(a) = \alpha$ and $F(b) = 1 - \alpha$. One then shows that $V(a, b)$ is maximized over all possible pairs (a, b) by the choice of a and b obtained by placing all the contaminating mass at $+\infty$.

2. Maximizing the asymptotic variance. Assume that ε and α are fixed, with $0 < \varepsilon < \alpha < \frac{1}{2}$. Let F_0 be a fixed distribution function with a density function $f_0 = F_0'$ satisfying the following two assumptions:

ASSUMPTION 1. f_0 is symmetric about 0, i.e., $f_0(x) = f_0(-x)$ a.e. x.

ASSUMPTION 2. $f_0(x)$ is strictly decreasing in $x > 0$.

The problem is to maximize $V(F)$, given by (1), over all F of form (3). Simplification of (1) yields

$$V(F) = V_1(F)/(1 - 2\alpha)^2, \tag{4}$$

where

$$V_1 = V_1(F) = -\left[c(\alpha)^2\right] + \int_a^b x^2\,dF + \alpha(a^2 + b^2). \tag{5}$$

Here $c(\alpha)$ is given by (2), and a and b satisfy

$$F(a) = \alpha, \qquad F(b) = 1 - \alpha. \tag{6}$$

Our first step will be to maximize V_1 subject to a and b being given. That is, we will maximize $V_1(F)$ over the convex subclass of distributions of form $F(x) = (1 - \varepsilon)F_0(x) + \varepsilon G(x)$ which satisfy (6). Only pairs of values of a and b for which this convex subclass is nonempty will be considered. For given values of a and b, $G(a)$ and $G(b)$ are determined by

$$\begin{aligned} G(a) &= \left(\alpha - (1 - \varepsilon)F_0(a)\right)/\varepsilon, \\ 1 - G(b) &= \left(\alpha - (1 - \varepsilon)(1 - F_0(b))\right)/\varepsilon, \end{aligned} \tag{7}$$

so that the G-mass $G(b) - G(a)$ is also known.

We may assume that a and b satisfy

$$0 \le |a| \le b, \tag{8}$$

so that $1 - F_0(b) \le F_0(a)$ and $1 - G(b) \ge G(a)$. The reason that there is no loss of generality in assuming that (8) holds is that if $F = (1 - \varepsilon)F_0 + \varepsilon G$ and if $F^* = (1 - \varepsilon)F_0 + \varepsilon G^*$, where $G^*(x) = 1 - G(-x)$ for all x, then clearly $V(F) = V(F^*)$ by symmetry.

For fixed a and b, it follows from (2), (3), and (5) that

$$V_1 = -\left[C_1 + \varepsilon\int_a^b x\,dG(x)\right]^2 + C_2 + \varepsilon\int_a^b x^2\,dG(x), \tag{9}$$

where C_1 and C_2 are positive constants, so that V_1 is a simple quadratic function of the moments

$$u = \int_a^b x\,dG(x), \qquad v = \int_a^b x^2\,dG(x). \tag{10}$$

It is well known that the pair $(u/\rho, v/\rho)$ can be any point in the convex hull of the curve $S = \{(x, x^2) \colon a \le x \le b\}$, where $\rho = G(b) - G(a)$ is fixed by (7). The upper boundary of the convex hull of S is a straight line segment, and each point on that line segment can be realized by a distribution whose restriction to the interval $[a, b]$ is supported by the pair of points $\{a, b\}$. Keeping u fixed and first maximizing V_1 relative to v, it is obvious from (9) that any distribution G maximizing V_1 must have v maximal, that is, must be supported by $\{a, b\}$.

For fixed a and b, let the maximizing G have masses p and q at a and b, respectively, so that $p + q = \rho$. Then we have $u = pa + qb = \rho a + q(b - a)$ and $v = pa^2 + qb^2 = \rho a^2 + q(b^2 - a^2)$. So the maximum value of V_1, subject to (6), is the maximum value of

$$\begin{aligned} V_1 = V_1(q) = &-\Big\{(1 - \varepsilon)\int_a^b x\,dF_0(x) + \varepsilon[(G(b) - G(a))a + q(b - a)] \\ &\qquad + \alpha(a + b)\Big\}^2 \\ &+ (1 - \varepsilon)\int_a^b x^2\,dF_0(x) + \varepsilon[(G(b) - G(a))a^2 + q(b^2 - a^2)] \\ &+ \alpha(a^2 + b^2), \end{aligned} \tag{11}$$

where q ranges over $[0, G(b) - G(a)]$, and where $G(a)$ and $G(b)$ are determined by (7).

Partial differention of (11) relative to q and substitution of identity (2) yields:

$$\begin{aligned} \partial V_1/\partial q &= -2c(\alpha)\varepsilon(b - a) + \varepsilon(b^2 - a^2) \\ &= 2\varepsilon(b - a)(1 - 2\alpha)\left[\frac{b + a}{2} - \frac{1}{1 - 2\alpha}\int_a^b x\,dF(x)\right], \end{aligned} \tag{12}$$

where $F = (1 - \varepsilon)F_0 + \varepsilon[(\rho - q)\delta_a + q\delta_b]$. Since we also have $\partial^2 V_1/\partial q^2 = -2\varepsilon^2(b - a) \le 0$, we need only inspect $\partial V_1/\partial q$ at 0 and ρ to determine whether V_1 is maximized when (i) $q = \rho$, (ii) $q \in (0, \rho)$, or (iii) $q = 0$. The following lemma shows that, when $\rho > 0$, the possibility of V_1 being maximized when $q = 0$ (corresponding to having all the contaminating mass at a) is ruled out under our assumptions.

LEMMA. *Let F_0 be a fixed distribution function satisfying Assumptions* 1 *and* 2. *Let a and b be fixed numbers which satisfy* (8) *and for which $\rho = G(b) - G(a)$ (defined by* (7)) *satisfies $\rho > 0$. Then $\partial V_1/\partial q$ is nonnegative at $q = 0$.*

PROOF. Suppose not, i.e., suppose that $\partial V_1/\partial q < 0$ at $q = 0$. By (12), this implies that when the restriction of the distribution to $[a, b]$ is $F = (1-\varepsilon)F_0 + \varepsilon\rho\delta_a$, the average of the distribution over $[a, b]$, namely $\int_a^b x\, dF(x)/(F(b) - F(a))$, is $> (b+a)/2$. Then the average over $[a, b]$ under F_0, namely $\int_a^b x\, dF_0(x)/(F_0(b) - F_0(a))$, must also be $> (b+a)/2$, since mixing F_0 with δ_a can only pull the average toward the left. Now let $x_0 = (b+a)/2$ and note that $x_0 \geq 0$ by assumption (8). To complete the proof by contradiction, it remains to show that the average value over $[a, b]$ under F_0 is $\leq x_0$. But this follows from the calculation

$$\begin{aligned}\int_a^b x\, dF_0(x) - x_0\int_a^b dF_0 &= \int_a^{x_0}(x - x_0)f_0(x)\, dx + \int_{x_0}^b (x - x_0)f_0(x)\, dx \\ &= \int_0^{(b-a)/2} t[f_0(x_0+t) - f_0(x_0 - t)]\, dt \leq 0,\end{aligned} \tag{13}$$

since Assumptions 1 and 2 imply that $f_0(x-t) \geq f_0(x+t)$ for all $t \geq 0$ and $x \geq 0$. □

In view of the lemma, the maximum of V_1 subject to fixed a and b satisfying (8) occurs when either (i) $q = \rho$ (all contaminating mass at b) or (ii) $q \in (0, \rho)$ (a proper mixture of mass at both a and b). We remark that calculations for the case when F_0 is the standard normal distribution show that both cases (i) and (ii) do in fact occur, depending on the values of a and b.

THEOREM. *Under Assumptions* 1 *and* 2 *on* F_0:

(i) *The maximum value of $V_1(F) = (1-2\alpha)^2 V(F)$ over all $F = (1-\varepsilon)F_0 + \varepsilon G$ is*

$$-\left[(1-\varepsilon)\int_a^b x\, dF_0(x) + \alpha(a+b)\right]^2 + (1-\varepsilon)\int_a^b x^2\, dF_0(x) + \alpha(a^2+b^2), \tag{14}$$

when $a = a_0$ and $b = b_0$, where $a_0 = F_0^{-1}(\alpha/(1-\varepsilon))$ and $b_0 = F_0^{-1}((1-\alpha)/(1-\varepsilon))$.

(ii) *The maximum is attained at $F = (1-\varepsilon)F_0 + \varepsilon G$ if and only if either G places mass* 1 *on $(b_0, \infty]$ or G places mass* 1 *on $[-\infty, -b_0)$.*

PROOF. (i) Let $V_1(a, b)$ denote the maximal value of V_1 subject to $F(a) = \alpha$ and $F(b) = 1 - \alpha$. We need to show $V_1(a, b) \leq V_1(a_0, b_0)$ for all possible (a, b). Without loss of generality, consider only pairs (a, b) which satisfy (8). Our first step is to show that for each *fixed* a, $V(a, b)$ is maximized at the maximal possible value of b (corresponding to $\rho = G(b) - G(a) = 0$); namely at $b = b(a)$ satisfying

$$F_0(b) = F_0(a) + (1 - 2\alpha)/(1 - \varepsilon). \tag{15}$$

For fixed a, we will show that $\partial V_1(a, b)/\partial b \geq 0$ for all b. First let S be any interval of bs for which $\partial V_1/\partial q$ (formula (12)) is ≥ 0 at $q = \rho$. Equivalently, S is an interval of bs for which

$$\frac{b + a}{2} - \frac{1}{1 - 2\alpha} \int_a^b x \, dF_b(x) \geq 0, \tag{16}$$

where F_b denotes a distribution with restriction to $[a, b]$ given by $F_b = (1 - \varepsilon)F_0 + \varepsilon\rho\delta_b$. In view of the lemma, $V_1(q)$ is maximized at $q = \rho$ for all $b \in S$. For $b \in S$, $V_1(a, b)$ is obtained by substituting $q = \rho = G(b) - G(a)$ into the right-hand side of (11). Differentiating $V_1(a, b)$ with respect to b, noting that $\partial(\varepsilon G(b))/\partial b = -(1 - \varepsilon)f_0(b)$ by (7), yields (after some simplification) that

$$\begin{aligned} \partial V_1(a, b)/\partial b &= 2[\varepsilon(G(b) - G(a)) + \alpha]\left[b - \left(\int_a^b x \, dF_b(x) + \alpha(a + b)\right)\right] \\ &= 2[\varepsilon(G(b) - G(a)) + \alpha] \\ &\quad \times\left[\left(b - \frac{b + a}{2}\right) + \left(\frac{b + a}{2}(1 - 2\alpha) - \int_a^b x \, dF_b(x)\right)\right] \geq 0 \end{aligned} \tag{17}$$

for all $b \in S$, by (16).

Next, for fixed a, let S_2 be any open interval of bs for which the value of q maximizing $V_1(q)$ in formula (11) satisfies $0 < q < \rho$. Then in view of the lemma, it follows that for $b \in S_2$, the maximum value of V_1 is

$$V_1(a, b) = -\left[\int_a^b x \, dF_q(x) + \alpha(a + b)\right]^2 + \int_a^b x^2 \, dF_q(x) + \alpha(a + b)^2, \tag{18}$$

where $F_q = (1 - \varepsilon)F_0 + \varepsilon(G(b) - G(a) - q)\delta_a + \varepsilon q\delta_b$ on $[a, b]$, and where $q = q(b)$ is the unique solution in $(0, \rho)$ of $\partial V(q)/\partial q = 0$. Equivalently, from (12), $q = q(b)$ satisfies

$$\frac{b + a}{2} - \frac{1}{1 - 2\alpha} \int_a^b x \, dF_q(x) = 0. \tag{19}$$

Differentiating (18) with respect to b on S_2 yields

$$\begin{aligned} \frac{\partial V_1(a, b)}{\partial b} &= -2\left[\int_a^b x \, dF_q(x) + \alpha(a + b)\right] \\ &\quad \times\left[(1 - \varepsilon)(b - a)f_0(b) + \varepsilon(b - a)\frac{dq}{db} + \varepsilon q + \alpha\right] \\ &\quad + (1 - \varepsilon)(b^2 - a^2)f_0(b) + \varepsilon(b^2 - a^2)\frac{dq}{db} + 2\varepsilon bq + 2\alpha b. \end{aligned} \tag{20}$$

Substitution of $(b + a)/2 = \int_a^b x \, dF_q(x) + \alpha(a + b)$ (which is just (19)) into (20) yields, after simplification, that

$$\frac{\partial V_1(a, b)}{\partial b} = (b - a)(\alpha + \varepsilon q) > 0 \tag{21}$$

for all $b \in S_2$.

From (17) and (21) it follows that, for each fixed a, $V_1(a, b)$ is maximized when b attains its maximum value, corresponding to $\rho = 0$. Thus $V_1(a)$, defined as the maximum value of V_1 given a, is given by (14) with $b = b(a)$ determined by (15). It remains to show that $V_1(a)$ attains its maximum at the maximum possible value of a, namely at $a = a_0 = F_0^{-1}(\alpha/(1 - \varepsilon))$.

Computation of $\partial V_1(a)/\partial a$, using from (15) that $\partial b/\partial a = f_0(a)/f_0(b)$, yields

$$\begin{aligned} \partial V_1(a)/\partial a = {} & -2\left[(1 - \varepsilon)\int_a^b x f_0(x)\,dx + \alpha(a + b)\right] \\ & \times\left[(1 - \varepsilon) f_0(a)(b - a) + \alpha(1 + f_0(a)/f_0(b))\right] \\ & + (1 - \varepsilon) f_0(a)(b^2 - a^2) + 2\alpha\left[a + b f_0(a)/f_0(b)\right]. \end{aligned} \tag{22}$$

So to show that $\partial V_1(a)/\partial a \geq 0$ for all a, completing the proof of (i), it suffices to show both:

$$b + a \geq 2\left[(1 - \varepsilon)\int_a^b x f_0(x)\,dx + \alpha(a + b)\right] \tag{23}$$

and

$$\frac{a + [b f_0(a)/f_0(b)]}{1 + [f_0(a)/f_0(b)]} \geq (1 - \varepsilon)\int_a^b x f_0(x)\,dx + \alpha(a + b). \tag{24}$$

Using an inequality from the proof of the lemma and using the identity (15) yields

$$\frac{b + a}{2} \geq \frac{\int_a^b x\,dF_0(x)}{F_0(b) - F_0(a)} = \frac{1 - \varepsilon}{1 - 2\alpha}\int_a^b x\,dF_0(x), \tag{25}$$

which is (23). Also (24) will follow from (23) if we can show that

$$2\frac{a f_0(b) + b f_0(a)}{f_0(a) + f_0(b)} \geq b + a. \tag{26}$$

But we have

$$2\frac{a f_0(b) + b f_0(a)}{f_0(a) + f_0(b)} - (b + a) = \frac{(b - a)(f_0(a) - f_0(b))}{f_0(a) + f_0(b)} \geq 0, \tag{27}$$

since $b - a > 0$ and since $b \geq |a|$ implies that $f_0(b) \leq f_0(|a|) = f_0(a)$ by Assumptions 1 and 2. This completes the proof of (i).

(ii) When $b > |a|$, it is easily seen that the inequalities (25), (26), and (27) are strict inequalities. Thus $\partial V(a)/\partial a > 0$ except at the boundary case where $|a| = b$, proving that the ***unique*** maximum value of V_1 is given by (14) when $a = a_0$ and $b = b_0$. Clearly $F = (1 - \varepsilon)F_0 + \varepsilon G$ attains the maximum subject to

(8) if and only if G is concentrated on $(b_0, \infty]$. Removal of side condition (8) completes the proof of (ii) by symmetry. □

REMARK 1. For numerical values of the maximal asymptotic variance corresponding to various values of ε and α when F_0 is the standard normal distribution, see Exhibit 4.9.2 on page 105 of Huber (1981).

REMARK 2. Not all F of the form $F = (1 - \varepsilon)F_0 + \varepsilon G$ satisfy the regularity conditions under which the α-trimmed mean is asymptotically normal with variance $V(F)$. For such regularity conditions, see Bickel (1965) and Stigler (1973). However it is clear from part (ii) of our theorem that there are suitably regular F's which attain the maximal value of V.

Acknowledgment. The author wishes to thank the referee for detailed suggestions leading to a much shorter and better presentation of the results. In particular the present proof of the lemma is due to the referee.

REFERENCES

ANDREWS, D. F., BICKEL, P. J., HAMPEL, F. R., HUBER, P. J., ROGERS, W. H. and TUKEY, J. W. (1972). *Robust Estimates of Location*. Princeton Univ. Press.
BICKEL, P. J. (1965). On some robust estimates of location. *Ann. Math. Statist.* **36** 847–858.
HUBER, P. J. (1981). *Robust Statistics*. Wiley, New York.
STIGLER, S. M. (1973). The asymptotic distribution of the trimmed mean. *Ann. Statist.* **1** 472–477.

DEPARTMENT OF MATHEMATICS
AND STATISTICS
UNIVERSITY OF CALGARY
CALGARY, ALBERTA
CANADA T2N 1N4