ELSEVIER

Statistics & Probability Letters 22 (1995) 79–85

STATISTICS & PROBABILITY LETTERS

On the variance of the trimmed mean

Philippe Capéraà*, Louis-Paul Rivest

Département de mathématiques et de statistique, Université Laval, Cité Universitaire, Québec, G1K7P4, Canada

Received November 1993; revised February 1994

Abstract

This note provides a new expression for the variance of the trimmed mean in terms of product moments of order statistics. An exact expression for the bias of the estimator of the variance of trimmed mean is also provided. As an illustration, exact calculations are carried out for small samples from various t-distributions using (Tiku and Kumra, 1985) product moment tables for the order statistics of the t-distribution.

Key words: Order statistics; t-distribution; Winsorized mean

1. Introduction and main results

The trimmed mean is a popular estimator for the location parameter of an outlier prone symmetric distribution. Most of the theory concerning trimmed means relies on asymptotic approximations (see Lehmann, 1983). Small sample calculations of the efficiency of trimmed means use either Monte Carlo simulations (Andrews et al., 1972) or exact results based on the theory of order statistics (Gastwirth and Cohen, 1970). The standard estimator for the variance of the trimmed mean, obtained through an asymptotic linearization, is based on the winsorized variance. Its small sample bias, as an estimator of the small sample variance of the trimmed mean, has not been investigated thoroughly. Huber (1981) showed this variance estimator to be a jackknife estimator. Thus, by Efron and Stein (1981), one may expect its bias to be positive. However, the magnitude of this bias has, as far as we know, never been calculated for any distribution. This paper hopes to fill this gap by presenting new exact formulae for the variance of the trimmed mean useful when the trimming percentage is small. An exact expression for the bias of the normalized winsorized variance as an estimator of the variance of the trimmed mean is also derived. The variance of some trimmed means and the biases of their linearization variance estimators are then evaluated explicitly in small normal and t samples.

Let $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ denote the order statistics of a sample of size n drawn from a distribution F which is symmetric with respect to 0. Distribution F is assumed to have a finite variance σ^2. David's (1981) notation is used throughout the paper: μ_i and μ_{ij} denote the moments and product moments of order

*Corresponding author.

SSDI 0167-7152(94)00052-A

statistics of a sample of size n: $\mu_i = E(X_{i:n})$ and $\mu_{ij} = E(X_{i:n}X_{j:n})$. The hypothesis that F is symmetric with respect to 0 implies that

$$\mu_{ij} = \mu_{(n-i+1)(n-j+1)} \quad \text{for } 1 \leqslant i, j \leqslant n. \tag{1.1}$$

For an integer k less than $n/2$, the k-trimmed mean $\hat{\theta}_{k:n}$ is defined as,

$$\hat{\theta}_{k:n} = \frac{1}{n-2k} \sum_{i=k+1}^{n-k} X_{i:n}.$$

Since F is symmetric, the variance of $\hat{\theta}_{k:n}$ can be expressed as

$$\operatorname{Var}(\hat{\theta}_{k:n}) = \frac{1}{(n-2k)^2} \sum_{j=k+1}^{n-k} \sum_{i=k+1}^{n-k} \mu_{ij}.$$

When the trimming fraction $2k/n$ is small, this expression for the variance of the trimmed mean is not evaluated easily, even in small samples, since it depends on several product moments of order statistics. For small k's, the following expression, whose derivation is presented in Section 2, is more suitable;

$$\operatorname{Var}(\hat{\theta}_{k:n}) = \frac{1}{(n-2k)^2}\left(n\sigma^2 - 2\sum_{i=1}^{k} \mu_{ii} + 4\sum_{i=1}^{k}\sum_{j=i+1}^{k+1} \mu_{ij} + 2\sum_{i=1}^{k}\sum_{j=n-k+1}^{n} \mu_{ij}\right). \tag{1.2}$$

This formula is quite simple for small k's. When k is equal to 1 one gets

$$\operatorname{Var}(\hat{\theta}_{1:n}) = \frac{n\sigma^2 - 2\mu_{11} + 4\mu_{12} + 2\mu_{1n}}{(n-2)^2},$$

when k is equal to 2,

$$\operatorname{Var}(\hat{\theta}_{2:n}) = \frac{n\sigma^2 - 2\mu_{11} - 2\mu_{22} + 4\mu_{12} + 4\mu_{13} + 4\mu_{23} + 2\mu_{1n} + 2\mu_{1n-1} + 2\mu_{2n-1} + 2\mu_{2n}}{(n-4)^2}.$$

The usual variance estimator for the trimmed mean is proportional to the winsorized variance (see Patel et al., 1988). It is given by

$$v(\hat{\theta}_{k:n}) = \frac{\left(\sum_{i=k+1}^{n-k} (X_{i:n} - \hat{\psi}_{k:n})^2 + k(X_{k+1:n} - \hat{\psi}_{k:n})^2 + k(X_{n-k:n} - \hat{\psi}_{k:n})^2\right)}{(n-2k)(n-2k-1)}, \tag{1.3}$$

where $\hat{\psi}_{k:n}$ is the winsorized mean,

$$\hat{\psi}_{k:n} = \frac{1}{n}\left\{\sum_{i=k+1}^{n-k} X_{i:n} + k(X_{n-k:n} + X_{k+1:n})\right\}.$$

An exact expression for the expectation of $v(\hat{\theta}_{k:n})$ is derived in Section 3. It is given by

$$E(v(\hat{\theta}_{k:n})) = \frac{1}{(n-2k)(n-2k-1)}\left(n\sigma^2 - 2\sum_{i=1}^{k} \mu_{ii} + 2\left(k - \frac{(k+1)^2}{n}\right)\mu_{(k+1)(k+1)} - 2\frac{(k+1)^2}{n}\mu_{(k+1)(n-k)}\right.$$
$$\left. - \frac{(n-2k-2)^2}{n}\operatorname{Var}(\hat{\theta}_{k+1:n}) + 4\frac{k+1}{n}\sum_{i=1}^{k+1}[\mu_{i(k+2)} + \mu_{(k+1)(n-i+1)}]\right). \tag{1.4}$$

Table 1
Exact and asymptotic efficiencies, with respect to the sample mean, of the 1-trimmed mean and of the 2-trimmed mean for two t-distributions

Sample size	Efficiencies of the 1-trimmed mean				Efficiencies of the 2-trimmed mean			
	t with df = 5		t with df = 3		t with df = 5		t with df = 3	
	exact	asymptotic	exact	asymptotic	exact	asymptotic	exact	asymptotic
7	1.1666	1.244	1.6837	1.936	1.1159	1.183	1.7314	1.948
8	1.1731	1.243	1.6825	1.912	1.1483	1.207	1.7742	1.969
9	1.1761	1.240	1.6766	1.888	1.1692	1.223	1.7974	1.975
10	1.1772	1.236	1.6684	1.865	1.1831	1.232	1.8092	1.974
11	1.1769	1.232	1.6591	1.843	1.1924	1.238	1.8139	1.968
12	1.1759	1.227	1.6492	1.822	1.1987	1.242	1.8141	1.959
13	1.1744	1.223	1.6392	1.803	1.2030	1.243	1.8115	1.948
14	1.1726	1.219	1.6293	1.785	1.2056	1.244	1.8069	1.936
15	1.1706	1.214	1.6196	1.768	1.2072	1.244	1.8010	1.924
16	1.1684	1.210	1.6101	1.753	1.2080	1.243	1.7943	1.912
17	1.1663	1.206	1.6010	1.738	1.2082	1.242	1.7870	1.900
18	1.1641	1.202	1.5923	1.724	1.2079	1.240	1.7795	1.888
19	1.1619	1.199	1.5839	1.711	1.2073	1.238	1.7717	1.876
20	1.1597	1.195	1.5758	1.699	1.2064	1.236	1.7640	1.865

The exact efficiencies of the $\hat{\theta}_{1:n}$ and $\hat{\theta}_{2:n}$ for the normal distribution and for t-distributions with 3 and 5 degrees of freedom were evaluated, using formula (1.2), in samples of sizes 7 to 20. The calculations were based on Teichroew tables for the product moments of order statistics from the normal distribution as reproduced in Sarhan and Greenberg (1962) and on Tiku and Kumra (1985) corresponding tables for the t-distribution. They were compared to the corresponding asymptotic efficiencies calculated using the formula for the asymptotic variance of the trimmed mean (Huber, 1981). For the normal distribution the two efficiencies were for all practical purposes equal; their relative differences were less than 1.5% for samples larger than 6. For the two t-distributions however, asymptotic approximations overestimate the efficiency of the trimmed mean. Asymptotic and exact efficiencies for t-distributions are compared in Table 1.

The biases of variance estimators $v(\hat{\theta}_{1:n})$ and $v(\hat{\theta}_{2:n})$ were also evaluated using (1.4); the results are not reported here. The relative biases of these variance estimators were all positive; they were less than 3% in samples of size larger than 7. Thus variance estimator $v(\hat{\theta}_{k:n})$ can be used safely even in very small samples. Trimmed t-statistics constructed with variance estimators similar to $v(\hat{\theta}_{k:n})$ are considered by Patel et al. (1988) and by Mudholkar et al. (1991). They propose accurate approximations to the null distributions of their studentized statistics and they show that trimmed t-statistics are more powerful than the classical Student's t for long tailed distributions.

2. Derivation of an exact expression for $\mathrm{Var}(\hat{\theta}_{k:n})$

The derivation is based on a conditional argument. Given the order statistics $X_{k:n}$ and $X_{n-k+1:n}$, the random variables $\{X_{k+1:n}, X_{k+2:n}, \ldots, X_{n-k:n}\}$ are distributed as the order statistics of a random sample from

distribution $G_{k:n}$ defined by

$$G_{k:n}(x) = \begin{cases} 0, & x \leqslant X_{k:n}, \\ \dfrac{F(x) - F(X_{k:n})}{F(X_{n-k+1:n}) - F(X_{k:n})}, & X_{k:n} < x \leqslant X_{n-k+1:n}, \\ 1, & x > X_{n-k+1:n}. \end{cases}$$

One has

$$\mathrm{Var}(\hat{\theta}_{k:n}) = E(\mathrm{Var}(\hat{\theta}_{k:n} | X_{k:n}, X_{n-k+1:n})) + \mathrm{Var}(E(\hat{\theta}_{k:n} | X_{k:n}, X_{n-k+1:n})). \tag{2.1}$$

The first term of the right-hand side of (2.1) is easily evaluated by noting that given $X_{k:n}$ and $X_{n-k+1:n}$, $\hat{\theta}_{k:n}$ is the mean of a sample of size n-$2k$ from distribution $G_{k:n}(x)$. An unbiased estimator of the conditional variance is therefore equal to

$$\frac{\sum_{i=k+1}^{n-k} (X_{i:n} - \hat{\theta}_{k:n})^2}{(n-2k)(n-2k-1)} = \frac{\sum_{i=k+1}^{n-k} X_{i:n}^2 - (n-2k)\hat{\theta}_{k:n}^2}{(n-2k)(n-2k-1)}.$$

Taking the expectation of this quantity yields

$$E(\mathrm{Var}(\hat{\theta}_{k:n} | X_{k:n}, X_{n-k+1:n})) = \frac{n\sigma^2 - 2\sum_{i=1}^{k} \mu_{ii}}{(n-2k)(n-2k-1)} - \frac{\mathrm{Var}(\hat{\theta}_{k:n})}{n-2k-1}, \tag{2.2}$$

since $E(\hat{\theta}_{k:n}) = 0$. To express the second term of (2.1) in terms of product moments of order statistics, note that

$$E(\hat{\theta}_{k:n} | X_{k:n}, X_{n-k+1:n}) = \int_{X_{k:n}}^{X_{n-k+1:n}} \frac{xf(x)}{F(X_{n-k+1:n}) - F(X_{k:n})} \mathrm{d}x, \tag{2.3}$$

where $f(x)$ is the density of F. By symmetry, the expectation of (2.3) is null. One has to calculate the expectation of the square of the right-hand side of (2.3) to complete the proof. Since the first moment of F is null, one has

$$E\left\{\left(\int_{X_{k:n}}^{X_{n-k+1:n}} \frac{xf(x)}{F(X_{n-k+1:n}) - F(X_{k:n})} \mathrm{d}x\right)^2\right\} = 2E\left\{\left(\int_{-\infty}^{X_{k:n}} \frac{xf(x)}{F(X_{n-k+1:n}) - F(X_{k:n})} \mathrm{d}x\right)^2\right\}$$
$$+ 2E\left(\int_{-\infty}^{X_{k:n}} \frac{xf(x)}{F(X_{n-k+1:n}) - F(X_{k:n})} \mathrm{d}x \int_{X_{n-k+1:n}}^{\infty} \frac{xf(x)}{F(X_{n-k+1:n}) - F(X_{k:n})} \mathrm{d}x\right). \tag{2.4}$$

Furthermore,

$$E\left\{\left(\int_{-\infty}^{X_{k:n}} \frac{xf(x)}{F(X_{n-k+1:n}) - F(X_{k:n})} \mathrm{d}x\right)^2\right\} = \frac{k(k+1)}{(n-2k)(n-2k-1)} E\left\{\left(\int_{-\infty}^{X_{k+2:n}} \frac{xf(x)}{F(X_{k+2:n})} \mathrm{d}x\right)^2\right\}. \tag{2.5}$$

This is proved by writing the left hand side of (2.5) in terms of the joint density of $X_{k:n}$ and $X_{n-k+1:n}$.

To evaluate the expectation on the right hand side of (2.5), one uses an argument similar to (2.1):

$$\operatorname{Var}\left(\sum_{i=1}^{k+1}\frac{X_{i:n}}{k+1}\right)=E\left(\operatorname{Var}\left(\sum_{i=1}^{k+1}\frac{X_{i:n}}{k+1}\middle|\, X_{k+2:n}\right)\right)$$
$$+E\left\{\left(\int_{-\infty}^{X_{k+2:n}}\frac{xf(x)}{F(X_{k+2:n})}\mathrm{d}x\right)^2\right\}-\left(\sum_{i=1}^{k+1}\frac{\mu_i}{k+1}\right)^2. \tag{2.6}$$

Given $X_{k+2:n}$, $\{X_{i:n}\}_1^{k+1}$ is an ordered sample of size $k+1$ from distribution $F(x)/F(X_{k+2:n})$ for $x \leqslant X_{k+2:n}$. Thus an argument similar to the one leading to (2.2) yields

$$E\left(\operatorname{Var}\left(\sum_{i=1}^{k+1}\frac{X_{i:n}}{k+1}\middle|\, X_{k+2:n}\right)\right)=\sum_{i=1}^{k+1}\frac{\mu_{ii}}{k(k+1)}-\frac{1}{k(k+1)^2}\left(\sum_{i=1}^{k+1}\mu_i\right)^2-\frac{1}{k}\operatorname{Var}\left(\sum_{i=1}^{k+1}\frac{X_{i:n}}{k+1}\right). \tag{2.7}$$

Putting this back in (2.6) and using (1.1) gives

$$E\left\{\left(\int_{-\infty}^{X_{k+2:n}}\frac{xf(x)}{F(X_{k+2:n})}\mathrm{d}x\right)^2\right\}=\frac{2}{k(k+1)}\sum_{i=1}^{k}\sum_{j=i+1}^{k+1}\mu_{ij}.$$

In a similar way one shows that

$$E\left(\frac{\int_{X_{n-k+1:n}}^{\infty}xf(x)\,\mathrm{d}x\int_{-\infty}^{X_{k:n}}xf(x)\,\mathrm{d}x}{(F(X_{n-k+1:n})-F(X_{k:n}))^2}\right)=\frac{1}{(n-2k)(n-2k-1)}\sum_{i=1}^{k}\sum_{j=n-k+1}^{n}\mu_{ij}.$$

Therefore it has been proved that

$$\operatorname{Var}(E(\hat{\theta}_{k:n}|X_{k:n},X_{n-k+1:n}))=\frac{2}{(n-2k)(n-2k-1)}\left(2\sum_{i=1}^{k}\sum_{j=i+1}^{k+1}\mu_{ij}+\sum_{i=1}^{k}\sum_{j=n-k+1}^{n}\mu_{ij}\right).$$

Combining this expression with (2.2) yields (1.2) as an expression for $\operatorname{Var}(\hat{\theta}_{k:n})$.

3. An expression for the expectation of $v(\hat{\theta}_{k:n})$

Variance estimator $v(\hat{\theta}_{k:n})$ can be written as

$$v(\hat{\theta}_{k:n})=\frac{\sum_{i=k+1}^{n-k}X_{i:n}^2+k(X_{k+1:n}^2+X_{n-k:n}^2)-n\hat{\psi}_{k:n}^2}{(n-2k)(n-2k-1)}.$$

Thus,

$$E(v(\hat{\theta}_{k:n}))=\frac{n\sigma^2-2\sum_{i=1}^{k}\mu_{ii}+2k\mu_{(k+1)(k+1)}-nE(\hat{\psi}_{k:n}^2)}{(n-2k)(n-2k-1)}.$$

The winsorized mean $\hat{\psi}_{k:n}$ is equal to

$$\hat{\psi}_{k:n} = \frac{(k+1)X_{k+1:n} + (k+1)X_{n-k:n} + (n-2k-2)\hat{\theta}_{k+1:n}}{n}.$$

Therefore

$$E(\hat{\psi}^2_{k:n}) = 2\left(\frac{k+1}{n}\right)^2 (\mu_{(k+1)(k+1)} + \mu_{(n-k)(k+1)}) + 2\left(\frac{n-2k-2}{n}\right)^2 \operatorname{Var}(\hat{\theta}_{k+1:n})$$
$$+ 4\frac{(k+1)(n-2k-2)}{n^2} E(X_{k+1:n}\hat{\theta}_{k+1:n}).$$

To complete the derivation, note that

$$(n-2k-2)E(X_{k+1:n}\hat{\theta}_{k+1:n}) = E\left(X_{k+1:n}\sum_{i=k+2}^{n} X_{i:n}\right) - \sum_{i=n-k}^{n} \mu_{(k+1)i}.$$

A conditional argument shows that

$$E\left(X_{k+1:n}\sum_{i=k+2}^{n} X_{i:n}\right) = (n-k-1)E\left(X_{k+1:n}\int_{X_{k+1:n}}^{\infty} \frac{xf(x)\,\mathrm{d}x}{(1-F(X_{k+1:n}))}\right)$$
$$= -(k+1)E\left(X_{k+2:n}\int_{-\infty}^{X_{k+2:n}} \frac{xf(x)\,\mathrm{d}x}{F(X_{k+2:n})}\right) = -\sum_{i=1}^{k+1} \mu_{i(k+2)}.$$

The final expression for $E(\hat{\psi}_{k:n^2})$ is then

$$E(\hat{\psi}^2_{k:n}) = 2\left(\frac{k+1}{n}\right)^2 (\mu_{(k+1)(k+1)} + \mu_{(n-k)(k+1)}) + 2\left(\frac{n-2k-2}{n}\right)^2 \operatorname{Var}(\hat{\theta}_{k+1:n})$$
$$- 4\frac{(k+1)}{n^2}\left(\sum_{i=1}^{k+1} \mu_{i(k+2)} + \sum_{i=n-k}^{n} \mu_{(k+1)i}\right).$$

Putting this in (3.1) yields expression (1.4) for the expectation of $v(\hat{\theta}_{k:n})$.

Acknowledgements

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada.

References

Andrews, D.F., P.J. Bickel, F.R. Hampel, P.J. Huber, W.H. Rogers and J.W. Tukey (1972), *Robust Estimates of Location. Survey and Advances* (Princeton University Press, Princeton, NJ).

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

Efron, B. and C. Stein (1981), The Jackknife estimate of variance, *Ann. Statist.* **9**, 586–596.

Gastwirth, J. and M. Cohen (1970), Small sample behavior of some robust linear estimates of location, *J. Amer. Statist. Assoc.* **65**, 946–976.

Huber, P.J. (1981), *Robust Statistics* (Wiley, New York).

Lehmann, E. (1983), *Theory of Point Estimation* (Wiley, New York).

Mudholkar, A., G.S. Mudholkar and D.K. Srivastava (1991), A construction and appraisal of pooled trimmed-t statistics, *Comm. Statist. Theory Methods* **20**, 1345–1359.

Patel, K.R., G.S. Mudholkar and J.L.I. Fernando (1988), Student t approximations for three simple robust estimators, *J. Amer. Statist. Assoc.* **83**, 1203–1211.

Sarhan, A.E. and B.G. Greenberg (1962), *Contributions to Order Statistics* (Wiley, New York).

Tiku, M.L. and S. Kumra (1985), Expected values and variances and covariances of order statistics for a family of symmetric distributions (Student's t), *Selected Tables in Mathematical Statistics* (American Mathematical Society, Providence, RI).