South African Statistical Journal

SUPPLEMENTARY MATERIAL FOR "ESTIMATION OF LOCATION PARAMETER WITHIN PRE-SPECIFIED ERROR BOUND WITH SECOND-ORDER EFFICIENT TWO-STAGE PROCEDURE"

Bhargab Chattopadhyay[1]
Department of Decision Sciences and Information Systems, Indian Institute of Management Visakhapatnam, Andhra Pradesh, India
e-mail: *bhargab@iimv.ac.in*

Swarnali Banerjee
Department of Mathematics and Statistics, Loyola University Chicago, Chicago, IL, USA
e-mail: *sbanerjee@luc.edu*

This supplement contains an illustration of a modified mean absolute deviation-based procedure for finding a fixed-width confidence interval for the normal mean.

Key words: Asymptotic efficiency, Taylor's theorem, Two-stage confidence interval procedure.

Herrey (1965) argued that when slight deviations from normality occur due to non-statistical reasons in data arising out of physical experiments, it is not unusual to assume normality. In this context, Chattopadhyay and Mukhopadhyay (2013) introduced a modified two-stage procedure for constructing fixed-width confidence interval for the normal population mean using mean absolute deviation (MAD) which enjoys only first order efficiency property. In this supplementary material, we provide an illustration of our modified two-stage procedure and a simulation study for constructing a fixed-width confidence interval for the mean of normal distribution under suspect outliers using Mean Absolute Deviation (MAD) as an estimator of population standard deviation. Please note that in this case, conditions (a)–(f) of the main paper are satisfied.

1. MAD-Based modified two-stage procedure

Suppose $X_1, \ldots.., X_m$ are i.i.d. normal random variables with population mean μ and standard deviation σ. Here $(\mu, \sigma^2) \in (\Re \times \Re^2)$. Let the sample mean based on $X_1, \ldots, X_m$, $\bar{X}_m$ and $T_m = T_m(X_1, \ldots, X_m)$ be unbiased estimators of μ and σ respectively. Here, T_m is an unbiased estimator of population standard deviation based on MAD such that

$$T_m = c_m^{-1} MAD_m, \text{ where, } MAD_m = (m)^{-1} \sum_{1 \le i \le m} \left| X_i - \overline{X} \right| \tag{1.1}$$

[1]Corresponding author.
MSC2020 subject classifications. 62F99, 62L12, 62L10.

and

$$c_m = \sqrt{\frac{2}{\pi}}\sqrt{\frac{m-1}{m}}. \tag{1.2}$$

Now, let W_m be the MAD-based standardized sample mean. Suppose that $\sigma > \sigma_L(> 0)$ and σ_L is known. Now, we know that $C > a^2\sigma_L^2/d^2$, a known lower bound. Along the lines of Mukhopadhyay and Duggan (1997), for $m_0 \geq 2$ we define:

$$m \equiv m(d) = \max\left\{m_0, \left\langle z_{\alpha/2}^2\sigma_L^2/d^2\right\rangle + 1\right\}, \tag{1.3}$$

We begin with pilot observations $X_1, ..., X_m$ and define the final sample size:

$$Q_{MAD} \equiv Q_{MAD}(d) = \max\left\{m, \left\langle b_{m,\alpha/2}^2 T_m^2/d^2\right\rangle + 1\right\}. \tag{1.4}$$

If $Q_{MAD} = m$, no further observations are collected beyond the pilot set, but if $Q_{MAD} > m$, then we collect $Q_{MAD} - m$ additional observations in the second stage. Finally, based on the combined data $X_1, ..., X_Q$ from both stages, we construct the following fixed-width confidence interval

$$J_{Q_{MAD}} = \left[\overline{X}_{Q_{MAD}} \pm d\right] \tag{1.5}$$

for μ. In the next section, we consider the properties enjoyed by our two-stage procedure from (1.4) and (1.5).

2. Characteristics

Before we consider the properties enjoyed by our two-stage procedure, we look at the lemma 1 which gives the approximate expression of the percentile point $b_{m,\alpha/2}$ in terms of $z_{\alpha/2}$ (upper $100(\alpha/2)\%$ points of the distribution of $N(0,1)$) and m (Pilot sample size).

Lemma 1. *Suppose that $b_{m,\alpha/2}$ and $z_{\alpha/2}$ are the upper $100(\alpha/2)\%$ points of the distribution of W_m and standard normal distribution respectively. Then the approximate expression of the percentile point $b_{m,\alpha/2}$ in terms of $z_{\alpha/2}$ and the sample size m is:*

$$b_{m,\alpha/2} = z_{\alpha/2} + \frac{b_1}{m_1} + O(m_1^{-2}), b_1 = B_{01} - B_{11} \tag{2.6}$$

Proof. The proof is given in the Appendix. ■

Theorem 1. *For the MAD-based two-stage procedure from (1.4)-(1.5), $(\mu,\sigma) \in \Re \times \Re^+, 0 < \sigma_L < \sigma$ and α, we have:*

- *(i)* $P_{\mu,\sigma}\left\{\mu \in J_{Q_{MAD}}\right\} \geq 1 - \alpha$ *, for any fixed d [Exact Consistency];*
- *(ii)* $Q_{MAD}/C \xrightarrow{P} 1$ *as* d $\to 0$;
- *(iii)* $P_{\mu,\sigma}\left\{\mu \in J_{Q_{MAD}}\right\} \to 1 - \alpha$ as d $\to 0$ *[Asymptotic Consistency];*
- *(iv)* $E_{\mu,\sigma}\left[Q_{MAD}/C\right] \to 1$ *as* d $\to 0$ *[First-Order Efficiency];*
- *(v)* $E_{\mu,\sigma}\left[Q_{MAD} - C\right]$ *is bounded as* d $\to 0$ *[Second-Order Efficiency].*

Proof. Parts (i)–(iv) is restating of Theorem 3.1 of Chattopadhyay and Mukhopadhyay (2013). So, we now prove part (v). Using (1.4), we write the following basic inequality:

$$\frac{b_{m,\alpha/2}^2 T_m^2}{d^2} \le Q_{MAD} \le mI(Q_{MAD} = m) + \frac{b_{m,\alpha/2}^2 T_m^2}{d^2} + 1. \tag{2.7}$$

Thus, we can write

$$\begin{aligned}\left(b_{m,\alpha/2}^2 - z_{\alpha/2}^2\right) E(T_m^2) d^{-2} \le E_{\mu,\sigma}\left[Q_{MAD} - C\right] \le mP_{\mu,\sigma}\left(Q_{MAD} = m\right)& \\ + \left(b_{m,\alpha/2}^2 - z_{\alpha/2}^2\right) E(T_m^2) d^{-2} + 1.& \end{aligned} \tag{2.8}$$

Before we proceed, let us consider lemma 2 which is related to the first term in the right hand side of the inequality defined in (2.8).

Lemma 2. *For the MAD-based two-stage procedure from (1.4)-(1.5), for all* $(\mu, \sigma) \in \mathfrak{R} \times \mathfrak{R}^+$, $0 < \sigma_L < \sigma$ *and* $0 < \alpha < 1$, *we have:*

$$P_{\mu,\sigma}\left(Q_{MAD} = m\right) = \frac{32}{m}\left(\frac{\sigma}{\sigma - \sigma_L}\right)^4 + O(m^{-2}). \tag{2.9}$$

Proof. Here, we proceed along the lines of Mukhopadhyay and Duggan (1997). For a small $d(> 0)$ note that:

$$P_{\mu,\sigma}\left(Q_{MAD} = m\right) = P_{\mu,\sigma}\left(T_m^2 < md^2/b_{m,\alpha/2}^2\right) = P_{\mu,\sigma}\left(\frac{T_m}{\sigma} - 1 < h_m\right), \tag{2.10}$$

where $h_m = \frac{m^{1/2} d}{\sigma b_{m,\alpha/2}} - 1$. But, we observe:

$$\begin{aligned} m^{1/2} d &= z_{\alpha/2}\sigma_L + o(1) \\ b_{m,\alpha/2} &= z_{\alpha/2} + o(1) \end{aligned} \tag{2.11}$$

so that we can claim:

$$h_m = \left(\sigma_L \sigma^{-1} - 1\right) + o(1). \tag{2.12}$$

Thus, for sufficiently large m, that is for sufficiently small $d(> 0)$, we may claim $h_m < \frac{1}{2}\left(\sigma_L \sigma^{-1} - 1\right)$ and this upper bound is negative. Using the expression of the moments of T_m, we have,

$$\begin{aligned} E_{\mu,\sigma}\left(\frac{T_m}{\sigma} - 1\right)^4 =& E\left(\frac{T_m}{\sigma}\right)^4 - 4E\left(\frac{T_m}{\sigma}\right)^3 + 6E\left(\frac{T_m}{\sigma}\right)^2 - 3 \\ =& \frac{2}{m} + O(m^{-2}) \end{aligned} \tag{2.13}$$

Recall that the upper bound for h_m is negative. Hence, from (2.10)-(2.13) we can conclude that for large m:

$$P_{\mu,\sigma}\left(Q_{MAD}=m\right)=P_{\mu,\sigma}\left(\frac{T_m}{\sigma}-1<h_m\right)\leq P_{\mu,\sigma}\left(\left|\frac{T_m}{\sigma}-1\right|>\frac{1}{2}\left(1-\frac{\sigma_L}{\sigma}\right)\right)$$
$$\leq\left\{\frac{1}{2}\left(1-\frac{\sigma_L}{\sigma}\right)\right\}^{-4}E_{\mu,\sigma}\left(\left|\frac{T_m}{\sigma}-1\right|^4\right)=\frac{32}{m}\left(\frac{\sigma}{\sigma-\sigma_L}\right)^4+O(m^{-2}).$$

This proves lemma 2. ■

Also, from lemma 1, we have:

$$b^2_{m,\alpha/2}z^{-2}_{\alpha/2}=1+2b_1(z_{\alpha/2}m)^{-1}+O\left(m^{-2}\right). \tag{2.14}$$

Finally, we combine (2.8), (2.9) and (2.14) and the expression of $E(T_m^2)$ to get:

$$2z_{\alpha/2}b_1\sigma^2\sigma_L^{-2}+o\left(d\right)\leq E_{\mu,\sigma}\left[Q_{MAD}-C\right]\leq 2z_{\alpha/2}b_1\sigma^2\sigma_L^{-2}+32\left(\frac{\sigma}{\sigma-\sigma_L}\right)^4+1+o\left(d\right). \tag{2.15}$$

Part (v) follows trivially from (2.15). ■

3. Real data illustration

This example indicates that the MAD-based test is more robust than the usual t-test. The data used in this article was first used in Welch (1987). This data came from an experiment conducted to test fault reduction method on telephone lines. The data consists of inverse test and control fault rates in 14 matched pairs of areas. To test if the inverse fault rate differences for telephone lines are negligible, Welch (1987) assumed that the data was a random sample drawn from a normal population with unknown mean μ and variance σ^2.

Table 1 shows the inverse fault-rate differences for telephone lines in 14 pairs of areas. The mean $\bar{X}_{14}$, standard deviation s_{14} and the estimate of the unbiased estimator based on MAD, T_{14} are respectively given as $\bar{X}_{14}=38.78571, s_{14}=321.8328, T_{14}=291.3516$. Now suppose, we assume that the population standard deviation of the inverse fault-rate differences for telephone lines is atleast 1.137282, that is, we treat these observations as our pilot data set with m =14.

Here, we want to construct a 95% confidence interval for the inverse fault-rate differences for telephone lines of width 4 units. While observing the data, one may be tempted to discard one or two possible outlying observations. However, one should refrain from doing so without consulting the experimenter, as they may add valuable information and also there is huge sampling cost involved in these experiments. Thus, in total we have 14 observations in our pilot sample. When $m=14$, 95 percentile point of the distribution of $W_m=\sqrt{m}\overline{X}_m/T_m$ and the t distribution are respectively given as $b_{14,0.05}=2.433584$ and $t_{14,0.05}=2.160369$. Thus, based on all these 14 differences of table 1, the corresponding 95% fixed-width (50 units, say) confidence interval using modified Stein's two-stage procedure given by Mukhopadhyay and Duggan (1997) would require at least 760 additional observations while based on MAD given in (1.4)–(1.5), we would need at least 586 more

Table 1. Inverse fault-rate differences for telephone lines ((Test-Control)$\times 10^5$).

Pair	**Difference**	**Pair**	**Difference**
1	−988	8	−135
2	309	9	110
3	269	10	93
4	228	11	83
5	204	12	−78
6	197	13	59
7	189	14	3

observations. Clearly, there is a gain in using MAD based two-stage procedure instead of a modified Stein's two-stage procedure in case of suspect outliers.

4. Appendix

4.1 Expansion of the $E(T_m^2)$

The following expressions were obtained from Geary (1936):

$$E(MAD_m^2) = a^2\left(1 + \sum_{j=1}^{5} \frac{a'_{2j}}{m_1^j}\right) + O(m^{-6}), \quad a'_{2j} = \frac{a''_{2j}}{a^2} \quad \text{and} \quad m_1 = m - 1,$$

where $a = \sqrt{2/\pi}$ and

$$a''_{21} = 1 - a^2, \quad a''_{22} = 1.5a^2 - 1, \quad a''_{23} = 1 - 1.5a^2, \quad a''_{24} = \frac{37}{24}a^2 - 1, \quad a''_{25} = 1 - \frac{37}{24}a^2.$$

Note that $T_m = c_m^{-1} MAD_m$. So,

$$E(T_m^2) = \frac{\pi}{2}\left(1 + \frac{1}{m_1}\right) a^2 \left(1 + \sum_{j=1}^{5} \frac{a'_{2j}}{{m_1}^j}\right) + O(m^{-6}) = \left(1 + \sum_{j=1}^{5} \frac{a_{2j}}{{m_1}^j}\right) + O(m^{-6}),$$

where

$$a_{21} = 1 + a'_{21}, \quad a_{2j} = a'_{2j} + a'_{\overline{2j-1}}, \quad \text{for } j = 2, 3, 4, 5. \tag{4.16}$$

4.2 Expansion of the E(T_m^3)

The following expressions were obtained from Geary (1936):

$$E(MAD_m^3) = a^3\left(1 + \sum_{j=1}^{5} \frac{a'_{3j}}{{m_1}^j} + O(m_1^{-6})\right), \quad a'_{3j} = \frac{a''_{3j}}{a^3},$$

where

a''_{31}	a''_{32}	a''_{33}	a''_{34}	a''_{35}
$3a - 3a^3$	$6.5a^3 - 4a$	$8a - 12.5a^3$	$\frac{151}{8}a^3 - 12a$	$16a - \frac{201}{8}a^3$

(4.17)

So, the third raw moment is,

$$E(T_m^3) = \left(1 + \sum_{j=1}^{5} \frac{a_{3j}}{{m_1}^j}\right) + O(m^{-6}),$$

where

$$\begin{aligned}
a_{31} &= \left(a'_{31} + \frac{3}{2}\right),\\
a_{32} &= \left(\frac{3}{2}a'_{31} + a'_{32} + \frac{3}{8}\right),\\
a_{33} &= \left(\frac{3}{8}a'_{31} + \frac{3}{2}a'_{32} + a'_{33} - \frac{1}{16}\right),\\
a_{34} &= \left(\frac{3}{8}a'_{32} - \frac{1}{16}a'_{31} + \frac{3}{2}a'_{33} + a'_{34} + \frac{3}{128}\right),\\
a_{35} &= \left(\frac{3}{128}a'_{31} - \frac{1}{16}a'_{32} + \frac{3}{8}a'_{33} + \frac{3}{2}a'_{34} + a'_{35} - \frac{3}{256}\right).
\end{aligned}$$

4.3 Expansion of the E(T_m^4)

The following expressions were obtained from Geary (1936):

$$E(MAD_m^4) = a^4\left(1 + \sum_{j=1}^{5} \frac{a'_{4j}}{m_1^j}\right) + O(m_1^{-6}),\, a'_{4j} = \frac{a''_{4j}}{a^4},$$

where

a''_{41}	a''_{42}	a''_{43}
$6a^2 - 6a^4$	$3 - 16a^2 + 20a^4$	$-6 + 45a^2 - 56a^4$

(4.18)

a''_{44}	a''_{45}
$15 - 108a^2 + 133a^4$	$-30 + \frac{845}{4}a^2 - 258a^4$

(4.19)

So, the fourth raw moment is given by

$$E(T_m^4) = \left(1 + \sum_{j=1}^{5} \frac{a_{4j}}{{m_1}^j}\right) + O(m^{-6}),$$

where

$$
\begin{aligned}
a_{41} &= \left(2 + a'_{41}\right), \\
a_{42} &= \left(1 + 2a'_{41} + a'_{42}\right), \\
a_{43} &= \left(a'_{41} + 2a'_{42} + a'_{43}\right), \\
a_{44} &= \left(a'_{42} + 2a'_{43} + a'_{44}\right), \\
a_{45} &= a'_{43} + 2a'_{44} + a'_{45}.
\end{aligned}
$$

4.4 Proof of Lemma 1

Proceeding in the same way as Mukhopadhyay and Chattopadhyay (2012), we have

$$
\begin{aligned}
\Phi\left(b_{m,\alpha/2}t\right) \approx& \Phi\left(z_{\alpha/2}t\right) + (b_{m,\alpha/2} - z_{\alpha/2}) \left.\frac{\partial \Phi(xt)}{\partial x}\right]_{x=z_{\alpha/2}} \\
=& \Phi\left(z_{\alpha/2}t\right) + (b_{m,\alpha/2} - z_{\alpha/2})t\phi(z_{\alpha/2}t) \\
=& I + (b_{m,\alpha/2} - z_{\alpha/2})tII. \qquad (4.20)
\end{aligned}
$$

We now focus on the terms I and II successively and as each term is a function of t, we expand them around $t = 1$ to write

$$
\begin{aligned}
I = \Phi\left(z_{\alpha/2}t\right) \approx& \Phi\left(z_{\alpha/2}\right) + (t-1)z_{\alpha/2}\phi(z_{\alpha/2}) - \frac{1}{2}(t-1)^2 z^3_{\alpha/2}\phi(z_{\alpha/2}) \\
&- \frac{1}{6}(t-1)^3 z^3_{\alpha/2}(1 - z^2_{\alpha/2})\phi(z_{\alpha/2}) + \frac{1}{24}(t-1)^4 z^5_{\alpha/2}(3 - z^2_{\alpha/2})\phi(z_{\alpha/2}) \qquad (4.21)
\end{aligned}
$$

and

$$
\begin{aligned}
II = \phi(z_{\alpha/2}t) \approx& \phi\left(z_{\alpha/2}\right) - (t-1)z^2_{\alpha/2}\phi\left(z_{\alpha/2}\right) + \frac{1}{2}(t-1)^2 \\
&\times \left(z^4_{\alpha/2} - z^2_{\alpha/2}\right)\phi\left(z_{\alpha/2}\right) + \frac{1}{6}(t-1)^3\left(-z^6_{\alpha/2} + 3z^4_{\alpha/2}\right)\phi\left(z_{\alpha/2}\right). \qquad (4.22)
\end{aligned}
$$

The following are provided for brevity and precision:

$$
\begin{aligned}
\frac{\partial \Phi(xt)}{\partial t} &= x\phi(xt) \\
\frac{\partial^2 \Phi(xt)}{\partial t^2} &= -x^3 t\phi(xt) \\
\frac{\partial^3 \Phi(xt)}{\partial t^3} &= -x^3\phi(xt)\left(1 - x^2t^2\right) \\
\frac{\partial^4 \Phi(xt)}{\partial t^4} &= -x^5\phi(xt)\left(-3t + x^2t^3\right).
\end{aligned}
$$

We exploit these expressions as needed. Next, by combining the terms I, $(b_{m,\alpha/2} - z_{\alpha/2})tII$, we obtain

$$\Phi\left(b_{m,\alpha/2}t\right) \approx \Phi\left(z_{\alpha/2}\right) + \phi(z_{\alpha/2})\left[z_{\alpha/2}\left(t-1\right) - \frac{1}{2}\left(t-1\right)^2 z_{\alpha/2}^3 + \frac{1}{6}\left(t-1\right)^3\left(z_{\alpha/2}^5 - z_{\alpha/2}^3\right)\right.$$
$$+ \frac{1}{24}\left(t-1\right)^4\left(3z_{\alpha/2}^5 - z_{\alpha/2}^7\right) + t(b_{m,\alpha/2} - z_{\alpha/2})\left\{1 - \left(t-1\right) z_{\alpha/2}^2\right.$$
$$\left.\left. + \frac{1}{2}\left(t-1\right)^2\left(z_{\alpha/2}^4 - z_{\alpha/2}^2\right) + \frac{1}{6}\left(t-1\right)^3\left(-z_{\alpha/2}^6 + 3z_{\alpha/2}^4\right)\right\}\right], \tag{4.23}$$

for all $t > 0$. Then, replacing t with the statistic T_m and then taking expectations throughout, we obtain

$$E\left[\Phi\left(b_{m,\alpha/2}T_m\right)\right]$$
$$\approx \Phi\left(z_{\alpha/2}\right) + \phi(z_{\alpha/2})E\left[z_{\alpha/2}\left(T_m-1\right) - \frac{1}{2}\left(T_m-1\right)^2 z_{\alpha/2}^3 + \frac{1}{6}\left(T_m-1\right)^3\left(z_{\alpha/2}^5 - z_{\alpha}^3\right)\right.$$
$$+\frac{1}{24}\left(T_m-1\right)^4\left(3z_{\alpha/2}^5 - z_{\alpha/2}^7\right) + T_m(b_{m,\alpha/2} - z_{\alpha/2})\left(1 - \left(T_m-1\right) z_{\alpha/2}^2\right.$$
$$\left.\left.+\frac{1}{2}\left(T_m-1\right)^2\left(z_{\alpha/2}^4 - z_{\alpha/2}^2\right) + \frac{1}{6}\left(T_m-1\right)^3\left(-z_{\alpha/2}^6 + 3z_{\alpha/2}^4\right)\right)\right]. \tag{4.24}$$

So, alternatively,

$$E\left[\Phi\left(b_{m,\alpha/2}T_m\right)\right]$$
$$\approx \Phi\left(z_{\alpha/2}\right) + \phi(z_{\alpha/2})E\left[\left(\frac{1}{8}z_{\alpha/2}^5 - \frac{1}{24}z_{\alpha/2}^7\right)T_m^4 + \left(-\frac{1}{3}z_{\alpha/2}^5 - \frac{1}{6}z_{\alpha/2}^3 + \frac{1}{6}z_{\alpha/2}^7\right)T_m^3\right.$$
$$+\left(\frac{1}{4}z_{\alpha/2}^5 - \frac{1}{4}z_{\alpha/2}^7\right)T_m^2 + \left(\frac{1}{2}z_{\alpha/2}^3 + z_{\alpha/2} + \frac{1}{6}z_{\alpha/2}^7\right)T_m - z_{\alpha/2} - \frac{1}{24}z_{\alpha/2}^5$$
$$-\frac{1}{3}z_{\alpha/2}^3 - \frac{1}{24}z_{\alpha/2}^7 + \left(b_{m,\alpha/2} - z_{\alpha/2}\right)\left(\left(-\frac{1}{6}z_{\alpha/2}^6 + \frac{1}{2}z_{\alpha/2}^4\right)T_m^4\right.$$
$$\left.\left.+\left(-z_{\alpha/2}^4 - \frac{1}{2}z_{\alpha/2}^2 + \frac{1}{2}z_{\alpha/2}^6\right)T_m^3 + \left(1 + \frac{1}{2}z_{\alpha/2}^2 + \frac{1}{2}z_{\alpha/2}^6\right)T_m + \left(-\frac{1}{2}z_{\alpha/2}^6 + \frac{1}{2}z_{\alpha/2}^4\right)T_m^2\right)\right]. \tag{4.25}$$

Now, we recall that $E\left[\Phi(b_{m,\alpha/2}T_m)\right]$ and $\Phi\left(z_{\alpha/2}\right)$ are both the same as to $1-\alpha/2$. Thus, the remaining terms from (4.25) may be approximated by zero which gives us the following equation:

$$B_0 + B_1 b_{m,\alpha/2} = 0,$$

where the coefficients B_0, B_1, themselves involving m, can be expressed as

$$\begin{aligned}
B_1 =&\left(-\frac{1}{6}z_{\alpha/2}^6-\frac{1}{2}z_{\alpha/2}^4\right)E\left[T_m^4\right]+\left(\frac{1}{2}z_{\alpha/2}^6+\frac{1}{2}z_{\alpha/2}^2\right)E\left[T_m^3\right]\\
&+\left(\frac{1}{2}z_{\alpha/2}^4-\frac{1}{2}z_{\alpha/2}^6\right)E\left[T_m^2\right]+\left(1+\frac{1}{2}z_{\alpha/2}^2+\frac{1}{6}z_{\alpha/2}^6\right);\\
B_0 =&\left(\frac{1}{8}z_{\alpha/2}^7-\frac{3}{8}z_{\alpha/2}^5\right)E\left[T_m^4\right]+\left(-\frac{1}{3}z_{\alpha/2}^7+\frac{2}{3}z_{\alpha/2}^5+\frac{1}{3}z_{\alpha/2}^3\right)E\left[T_m^3\right]\\
&+\left(\frac{1}{4}z_{\alpha/2}^7-\frac{1}{4}z_{\alpha/2}^5\right)E\left[T_m^2\right]-\frac{3}{8}z_{\alpha/2}^7-\frac{1}{24}z_{\alpha/2}^5-\frac{1}{3}z_{\alpha/2}^3-z_{\alpha/2}.
\end{aligned} \tag{4.26}$$

Let us suppose that for each $j=0,1$, we express (4.26) as $B_j=B_{j0}\left(1+\sum_{i=1}^5\frac{B_{ji}}{m_1^i}\right)$, where

$$B_{10}=\frac{1}{3}z_{\alpha/2}^6+1 \text{ and } B_{00}=-z_{\alpha/2}\left(\frac{1}{3}z_{\alpha/2}^6+1\right). \tag{4.27}$$

For $i=1,\dots,5$,

$$\begin{aligned}
B_{1i}&=\frac{1}{B_{10}}\left[\left(-\frac{1}{6}z_{\alpha/2}^6+\frac{1}{2}z_{\alpha/2}^4\right)a_{4i}+\left(-z_{\alpha/2}^4-\frac{1}{2}z_{\alpha/2}^2+\frac{1}{2}z_{\alpha/2}^6\right)a_{3i}+\left(\frac{1}{2}z_{\alpha/2}^4-\frac{1}{2}z_{\alpha/2}^6\right)a_{2i}\right]\\
B_{0i}&=\frac{1}{B_{00}}\left[\left(\frac{1}{8}z_{\alpha/2}^7-\frac{3}{8}z_{\alpha/2}^5\right)a_{4i}+\left(-\frac{1}{3}z_{\alpha/2}^7+\frac{2}{3}z_{\alpha/2}^5+\frac{1}{3}z_{\alpha/2}^3\right)a_{3i}+\left(\frac{1}{4}z_{\alpha/2}^7-\frac{1}{4}z_{\alpha/2}^5\right)a_{2i}\right].
\end{aligned} \tag{4.28}$$

Thus, to evaluate the terms B_0, B_1, we will need the first four moments of T_m. These are given by,

$$E(T_m)=1\text{and } E(T_m^j)=1+\sum_{i=1}^5\frac{a_{ji}}{m_1^i}+O(m_1^{-6}),\, j=2,3,4. \tag{4.29}$$

where a_{1i}, a_{3i}, a_{4i}, $(i=1,..,5)$ are given in subsections 4.2, 4.3.

Using the moments in (4.29) and then using such updated expressions of B_0 and B_1, it is evident that these depend upon arbitrary negative powers of m. Thus proceeding along the lines detailed in section 1, the required percentile points of the distribution of W_m are as follows:

$$b_{m,\alpha/2}=z_{\alpha/2}+\frac{b_1}{m_1}+\frac{b_2}{m_1^2}+\frac{b_3}{m_1^3}+\frac{b_4}{m_1^4}+\frac{b_5}{m_1^5}+O(m_1^{-6}) \tag{4.30}$$

with the requisite coefficients

$$
\begin{aligned}
b_1 =& B_{01} - B_{11}\\
b_2 =& - B_{01}B_{11} - B_{12} + B_{11}^2 + B_{02}\\
b_3 =& B_{01}\left(-B_{12} + B_{11}^2\right) - B_{13} + 2B_{12}B_{11} - B_{02}B_{11} + B_{03}\\
b_4 =& 2B_{13}B_{11} + B_{12}^2 - B_{14} + B_{01}\left(-B_{13} + 2B_{12}B_{11}\right) +\\
& B_{04} + B_{02}\left(-B_{12} + B_{11}^2\right) - B_{03}B_{11}\\
b_5 =& B_{01}\left(2B_{13}B_{11} + B_{12}^2 - B_{14}\right) + B_{02}\left(-B_{13} + 2B_{12}B_{11}\right) -\\
& B_{15} + 2B_{14}B_{11} + 2B_{13}B_{12} - B_{11}^5 - B_{13}B_{11}^2 + B_{12}^2B_{11} -\\
& B_{11}\left(2B_{13}B_{11} + B_{12}^2\right) + B_{05} - B_{04}B_{11} + B_{03}\left(-B_{12} + B_{11}^2\right)
\end{aligned}
\tag{4.31}
$$

This is an approximate expression of the percentile point $b_{m,\alpha/2}$ in terms of $z_{\alpha/2}$ and the sample size m.

References

Chattopadhyay, B. and Mukhopadhyay, N. (2013). Two-stage fixed-width confidence intervals for a normal mean in the presence of suspect outliers. *Sequential Analysis*, **32**, 134–157.

Geary, R. C. (1936). Moments of the ratio of the mean deviation to the standard deviation for normal samples. *Biometrika*, **28**, 295–307.

Herrey, E. M. J. (1965). Confidence intervals based on the mean absolute deviation of a normal sample. *Journal of American Statistical Association*, **60**, 257–269.

Mukhopadhyay, N. and Chattopadhyay, B. (2012). Asymptotic expansion of percentiles for a sample mean standardized by GMD: An application in the normal case. *Journal of Japan Statistical Society*, **42**, 165–184.

Mukhopadhyay, N. and Duggan, W. T. (1997). Can a two-stage procedure enjoy second-order properties? *Sankhya, Series A*, **59**, 435–448.

Welch, W. J. (1987). Rerandomizing the median in matched-pairs designs. *Biometrika*, **74**, 609–614.