ELSEVIER

Statistics & Probability Letters 42 (1999) 185–192

STATISTICS & PROBABILITY LETTERS

Third-order expansion of mean squared error of medians ☆

J.S. Huang[a,b,*]

[a] Department of Mathematics and Statistics, University of Guelph, Guelph, Ont., Canada N1G 2W1
[b] Tunghai University, Taichung, Taiwan, ROC

Received October 1997

Abstract

Mean squared error (mse) of the sample median is calculated up to the third-order term. The expansion serves two purposes. One is to provide a rather accurate approximation to mse. The other is to contradict a time-honored statistical folklore which says "It never pays to base the median on an odd number of observations". © 1999 Elsevier Science B.V. All rights reserved

MSC: primary 62E20; secondary 62G30

Keywords: Median; Odd–even; Mean square error; Asymptotic; Nonparametric

1. Introduction

Let M_n be the median of a sample of size n. By calculating $\mathrm{var}(M_n)$ to the $\mathrm{O}(n^{-2})$ term, Hodges and Lehmann (1967) showed that the median based on an even number $n = 2m$ of observations is just as good as that based on the odd number $n = 2m+1$, up to the precision of their calculation. Specifically, they showed that for a population with a *symmetric* density f satisfying some regularity conditions,

$$\mathrm{var}(M_n) = \frac{1}{8f^2(0)}m^{-1} - \frac{f''(0) + 12f^3(0)}{64f^5(0)}m^{-2} + \mathrm{o}(m^{-2}) \tag{1}$$

for both even n $(=2m)$ and odd n $(=2m+1)$. It was noted that

- To the accuracy of this approximation, one should not use the median based on an odd number of observations since the median based on the next smaller even number is equally accurate.

It has since been implied in the statistical literature that it is not worth having an extra observation when computing the median from $2m$ observations:

☆ Supported by a Canadian NSERC grant. Work completed while the author was on leave at Tunghai University, Taichung, Taiwan.

* E-mail: jhuang@msnet.mathstat.uoguelph.ca.

PII: S0167-7152(98)00211-9

- "... when computing the median of an odd sample size one may as well discard an observation at random! ..." (Stigler, 1977, p. 544),
- "... thus, odd as it sounds, an additional observation when $n=2m$ results in a loss in the efficiency ..." (Cabrera et al., 1994, p. 350),
- "... increasing the sample size results in a loss of precision! ..." (Oosterhoff, 1994, p. 401),
- "... the precision of the median of a normal sample decreases as the even sample size $n=2m$ increases to $n=2m+1$ ..." (Mudholkar and Hutson, 1997, p. 266),

This even–odd phenomenon, often called "an oddity, an anomaly", has come to be accepted as a time-honored common wisdom. Recently, Cabrera et al. (1994) carried it one step further. They reported that "The real truth is much worse than what Hodges and Lehmann noted", that *symmetric or otherwise*, M_{2m} has strictly *lower* mean squared error (mse) than M_{2m+1}

$$\operatorname{mse}(M_{2m}) - \operatorname{mse}(M_{2m+1}) = -\frac{1}{16f^2(M)}m^{-3} + o(m^{-3}), \tag{2}$$

where f is the density of the population and M its median. Thus asymptotically the weak inequality

$$\operatorname{mse}(M_{2m}) \leqslant \operatorname{mse}(M_{2m+1}), \quad m=1,2,\ldots \tag{3}$$

is upgraded to strict inequality, and more importantly, that the "oddity" holds even for asymmetric distributions.

In this note the third-order expansion of the mean squared error of the median will be derived. The expansion serves two purposes. One is to provide a rather accurate approximation to the mse. The other is to show that mse does not typically increase when the sample size goes from $2m$ to $2m+1$.

2. A Counterexample

Let $X_1, X_2, \ldots, X_n$ be a sample from the exponential density $f(x)=e^{-x}$, $x>0$, and let the order statistics be denoted by $X_{1,n} \leqslant X_{2,n} \leqslant \cdots \leqslant X_{n,n}$. The exact values of the mean and the variance of exponential order statistics are well known (see David, 1981, p. 49). Simple calculation leads to

$$E(M_{2m+1}) = \sum_{i=m+1}^{2m+1} i^{-1}, \quad \operatorname{var}(M_{2m+1}) = \sum_{i=m+1}^{2m+1} i^{-2},$$

$$E(M_{2m}) = \sum_{i=m+1}^{2m} i^{-1} + \frac{1}{2m}, \quad \operatorname{var}(M_{2m}) = \sum_{i=m+1}^{2m} i^{-2} + \frac{1}{4m^2}.$$

It therefore follows that $\operatorname{var}(M_n)$ is monotonely decreasing in n,

$$\operatorname{var}(M_{2m}) - \operatorname{var}(M_{2m+1}) = \frac{1}{4m^2} - \frac{1}{(2m+1)^2} = \frac{4m+1}{4m^2(2m+1)^2} > 0.$$

So is the bias, $\operatorname{bias}(M_n) \equiv E(M_n) - \ln 2$,

$$\operatorname{bias}(M_{2m}) - \operatorname{bias}(M_{2m+1}) = \frac{1}{2m} - \frac{1}{2m+1} = \frac{1}{2m(2m+1)} > 0;$$

and so is the square of the bias. Thus, for the exponential distribution, the exact value of $\operatorname{mse}(M_n)$ is monotonely decreasing, not just for large n, but for all $n \geqslant 1$. Each additional observation continues to improve the precision of M_n, regardless of n being even or odd. This contradicts (2). See Table 2 for some numerical values.

3. Asymptotic expansion

In this section mse(M_n) will be calculated up to terms of order $O(m^{-3})$ for sample sizes even ($n=2m$) and odd ($n=2m+1$). Following Cabrera et al. (1994), the parent distribution F will be assumed to have a density f satisfying some smoothness conditions, but otherwise not necessarily symmetric.

Theorem 1. *For any density f which satisfies* (i) $f(M)>0$, (ii) *the fifth derivative* $f^{(5)}$ *is continuous in a neighborhood of M, and* (iii) $\int |x|^\delta f(x)\,\mathrm{d}x<\infty$ *for some* $\delta>0$,

$$\mathrm{bias}(M_{2m})=\tfrac{1}{8}c_2m^{-1}+(\tfrac{3}{64}c_4-\tfrac{1}{16}c_2)m^{-2}+(\tfrac{1}{32}c_2-\tfrac{3}{32}c_4+\tfrac{15}{512}c_6)m^{-3}+o(m^{-3}), \tag{4}$$

$$\mathrm{bias}(M_{2m+1})=\tfrac{1}{8}c_2m^{-1}+(\tfrac{3}{64}c_4-\tfrac{3}{16}c_2)m^{-2}+(\tfrac{9}{32}c_2-\tfrac{3}{16}c_4+\tfrac{15}{512}c_6)m^{-3}+o(m^{-3}), \tag{5}$$

$$\mathrm{bias}^2(M_{2m})-\mathrm{bias}^2(M_{2m+1})=\tfrac{1}{32}c_2^2m^{-3}+o(m^{-3}), \tag{6}$$

$$\begin{aligned}\mathrm{mse}(M_{2m})=&(\tfrac{1}{8}c_1^2)m^{-1}+(-\tfrac{3}{16}c_1^2+\tfrac{3}{32}c_1c_3+\tfrac{3}{64}c_2^2)m^{-2}\\&+(\tfrac{7}{32}c_1^2-\tfrac{5}{32}c_2^2+\tfrac{15}{512}c_3^2-\tfrac{9}{32}c_1c_3+\tfrac{15}{256}c_2c_4+\tfrac{15}{256}c_1c_5)m^{-3}+o(m^{-3}),\end{aligned} \tag{7}$$

$$\begin{aligned}\mathrm{mse}(M_{2m+1})=&(\tfrac{1}{8}c_1^2)m^{-1}+(-\tfrac{3}{16}c_1^2+\tfrac{3}{32}c_1c_3+\tfrac{3}{64}c_2^2)m^{-2}\\&+(\tfrac{9}{32}c_1^2-\tfrac{3}{16}c_2^2+\tfrac{15}{512}c_3^2-\tfrac{3}{8}c_1c_3+\tfrac{15}{256}c_2c_4+\tfrac{15}{256}c_1c_5)m^{-3}+o(m^{-3}),\end{aligned} \tag{8}$$

$$\mathrm{mse}(M_{2m})-\mathrm{mse}(M_{2m+1})=(\tfrac{3}{32}c_1c_3-\tfrac{1}{16}c_1^2+\tfrac{1}{32}c_2^2)m^{-3}+o(m^{-3}), \tag{9}$$

where

$$c_k=(k!)^{-1}\left(\frac{\mathrm{d}^k}{\mathrm{d}t^k}F^{-1}(t)\bigg|_{t=1/2}\right),\quad k=1,\ldots,6.$$

Proof. Consider first the case of odd sample size n ($=2m+1$). Proceeding as in Cabrera et al., we expand mse(M_n) around $U_{m+1,\,2m+1}=\frac{1}{2}$, where $M_n=X_{m+1,\,2m+1}\overset{\mathrm{d}}{=}F^{-1}(U_{m+1,\,2m+1})$, and $U_{m+1,\,2m+1}$ is the corresponding order statistic from the uniform distribution. Note that (see David, 1981, p. 36)

$$E(U^i_{m+1,\,2m+1})=\frac{m+1}{n+1}\frac{m+2}{n+2}\cdots\frac{m+i}{n+i}.$$

Writing $A\equiv U_{m+1,\,2m+1}-\frac{1}{2}$, we see that after some calculations,

$$E(A)=E(A^3)=E(A^5)=0,\quad E(A^2)=\frac{1}{4(2m+3)},$$

$$E(A^4)=\frac{3}{16(2m+3)(2m+5)},\quad E(A^6)=\frac{15}{64(2m+3)(2m+5)(2m+7)}$$

and that $E(A^r)=o(m^{-3})$, $r>6$. Keeping in mind the order of magnitude, we discard all the $o(m^{-3})$ order terms at each step in the following expansions:

$$\begin{aligned}\mathrm{bias}(M_{2m+1})&=E[F^{-1}(U_{m+1,\,2m+1})-M]\\&=E(c_1A+c_2A^2+\cdots+c_6A^6)+o(m^{-3})\end{aligned}$$

$$= \frac{64c_2m^2 + 24(16c_2 + c_4)m + 560c_2 + 84c_4 + 15c_6}{64(2m+3)(2m+5)(2m+7)} + \mathrm{o}(m^{-3})$$

$$= \frac{1}{8}c_2m^{-1} + \left(\frac{3}{64}c_4 - \frac{3}{16}c_2\right)m^{-2} + \left(\frac{9}{32}c_2 - \frac{3}{16}c_4 + \frac{15}{512}c_6\right)m^{-3} + \mathrm{o}(m^{-3}), \tag{10}$$

$$\begin{aligned}
\mathrm{mse}(M_{2m+1}) &= E[(F^{-1}(U_{m+1,2m+1}) - M)^2] \\
&= E[(c_1A + c_2A^2 + c_3A^3 + c_4A^4 + c_5A^5)^2] + \mathrm{o}(m^{-3}) \\
&= E[c_1^2A^2 + 2c_1c_2A^3 + (2c_1c_3 + c_2^2)A^4 + (2c_1c_4 + 2c_2c_3)A^5 \\
&\quad + (2c_1c_5 + c_3^2 + 2c_2c_4)A^6] + \mathrm{o}(m^{-3}) \\
&= \frac{1}{4}c_1^2(2m+3)^{-1} + \frac{3}{16}(c_2^2 + 2c_1c_3)(2m+3)^{-1}(2m+5)^{-1} \\
&\quad + \frac{15}{64}(2c_1c_5 + c_3^2 + 2c_2c_4)(2m+3)^{-1}(2m+5)^{-1}(2m+7)^{-1} + \mathrm{o}(m^{-3}) \\
&= \left(\frac{1}{8}c_1^2\right)m^{-1} + \left(-\frac{3}{16}c_1^2 + \frac{3}{32}c_1c_3 + \frac{3}{64}c_2^2\right)m^{-2} \\
&\quad + \left(\frac{9}{32}c_1^2 - \frac{3}{16}c_2^2 + \frac{15}{512}c_3^2 - \frac{3}{8}c_1c_3 + \frac{15}{256}c_1c_5 + \frac{15}{256}c_2c_4\right)m^{-3} + \mathrm{o}(m^{-3}).
\end{aligned}$$

The even n $(=2m)$ case is more complicated since M_{2m} is not a single-order statistic. Unlike the previous case, $F(M_n)$ is no longer distributed like its counterpart from the uniform $(0,1)$ distribution. Writing

$$\mathrm{bias}(M_{2m}) = E(M_{2m} - M) = \tfrac{1}{2}E[(X_{m,2m} - M) + (X_{m+1,2m} - M)], \tag{11}$$

we expand $X_{m,2m}$ and $X_{m+1,2m}$ each around $\frac{1}{2}$ in powers of $A \equiv U_{m,2m} - \frac{1}{2}$ and $B \equiv U_{m+1,2m} - \frac{1}{2}$, respectively. Noting that $E(A^iB^j)$ is of order $\mathrm{O}(m^{-[(i+j+1)/2]})$, we obtain from (11),

$$2\mathrm{bias}(M_{2m}) = E[(c_1A + \cdots + c_6A^6) + (c_1B + \cdots + c_6B^6)] + \mathrm{o}(m^{-3}), \tag{12}$$

which leads to (4) after substituting in the moments $E(A^i), E(B^i)$. The calculation of $\mathrm{mse}(M_{2m})$ is much more laborious due to the presence of cross-product terms:

$$\mathrm{mse}(M_{2m}) = E[(M_{2m} - M)^2] = \tfrac{1}{4}E[((X_{m,2m} - M) + (X_{m+1,2m} - M))^2]. \tag{13}$$

$$4\mathrm{mse}(M_{2m}) = E[((c_1A + \cdots + c_5A^5) + (c_1B + \cdots + c_5B^5))^2] + \mathrm{o}(m^{-3}). \tag{14}$$

After expanding the square and discarding higher-order terms ($A^iB^j, i+j > 6$), the quantity inside the square bracket above reduces to

$$\begin{aligned}
&(2c_2c_4 + c_3^2 + 2c_1c_5)A^6 + (2c_1c_5B + 2c_2c_3 + 2c_1c_4)A^5 \\
&\quad + (2c_2c_4B^2 + 2c_1c_4B + c_2^2 + 2c_1c_3)A^4 + (2c_3^2B^3 + 2c_2c_3B^2 + 2c_1c_3B + 2c_1c_2)A^3 \\
&\quad + (2c_2c_4B^4 + 2c_2c_3B^3 + 2c_2^2B^2 + 2c_1c_2B + c_1^2)A^2 \\
&\quad + (2c_1c_5B^5 + 2c_1c_4B^4 + 2c_1c_3B^3 + 2c_1c_2B^2 + 2c_1^2B)A \\
&\quad + (2c_2c_4 + c_3^2 + 2c_1c_5)B^6 + (2c_1c_4 + 2c_2c_3)B^5 + (2c_1c_3 + c_2^2)B^4 + 2c_1c_2B^3 + c_1^2B^2.
\end{aligned}$$

Taking expectation of the expression above then completes the proof, a task which would have taken weeks of calculation by hand without the help of symbolic languages such as Maple or Mathematica. It involves evaluating all product moments $E(A^iB^j)$, $i+j\leqslant 6$, substituting these moments into the expression above, and finally, discarding higher-order terms after collecting and sorting the terms according to powers of m^{-1}. Omitting all the messy details, we present only the final result:

$$\begin{aligned}\mathrm{mse}(M_{2m}) &= K((2m+1)(m+1)(2m+3)(m+2)(2m+5)(m+3))^{-1}+\mathrm{o}(m^{-3})\\ &= (\tfrac{1}{8}c_1^2)m^{-1}+(-\tfrac{3}{16}c_1^2+\tfrac{3}{64}c_2^2+\tfrac{3}{32}c_1c_3)m^{-2}\\ &\quad +(\tfrac{7}{32}c_1^2-\tfrac{5}{32}c_2^2+\tfrac{15}{512}c_3^2-\tfrac{9}{32}c_1c_3+\tfrac{15}{256}c_2c_4+\tfrac{15}{256}c_1c_5)m^{-3}+\mathrm{o}(m^{-3}),\end{aligned}$$

where

$$\begin{aligned}K \equiv{}& c_1^2m^5+(9c_1^2+\tfrac{3}{8}c_2^2+\tfrac{3}{4}c_1c_3)m^4\\ &+(\tfrac{119}{4}c_1^2+\tfrac{43}{16}c_2^2+\tfrac{15}{64}c_3^2+\tfrac{45}{8}c_1c_3+\tfrac{15}{32}c_1c_5+\tfrac{15}{32}c_2c_4)m^3\\ &+(\tfrac{171}{4}c_1^2+7c_2^2+\tfrac{63}{64}c_3^2+\tfrac{111}{8}c_1c_3+\tfrac{75}{32}c_1c_5+\tfrac{33}{16}c_2c_4)m^2\\ &+(\tfrac{45}{2}c_1^2+\tfrac{141}{16}c_2^2+\tfrac{51}{32}c_3^2+\tfrac{45}{4}c_1c_3+\tfrac{45}{16}c_1c_5+\tfrac{93}{32}c_2c_4)m\\ &+\tfrac{45}{16}c_2(c_4+2c_2). \qquad \square\end{aligned}$$

Corollary 1. *If F is symmetric (w.r.t. zero), then $f'(0)=c_2=0,\ c_1=1/f(0),\ c_3=-f''(0)/6f^4(0)$, and*

$$\begin{aligned}\mathrm{mse}(M_{2m})-\mathrm{mse}(M_{2m+1}) &= \tfrac{1}{32}c_1(3c_3-2c_1)m^{-3}+\mathrm{o}(m^{-3})\\ &= -\frac{1}{64f^5(0)}(f''(0)+4f^3(0))m^{-3}+\mathrm{o}(m^{-3}). \qquad (15)\end{aligned}$$

Notice that our results (7) and (8), specialized to symmetric f, are in agreement with (1). Our (9), however, contradicts (2). The question regarding (3) can now be answered by (9), which is seen to depend on the distribution only up to the first two derivatives of the density (at M). Contrary to the common wisdom, the coefficient in (9) may or may not be negative. See Section 4 for some examples.

From (6) it is clear that bias^2 always decreases with increasing n, regardless of parity. Together with (9) it therefore implies that it is the variance that is making the difference

$$\mathrm{var}(M_{2m})-\mathrm{var}(M_{2m+1})=(\tfrac{3}{32}c_1c_3-\tfrac{1}{16}c_1^2)m^{-3}+\mathrm{o}(m^{-3}). \qquad (16)$$

When the increase in variance more than compensates for the decrease in bias^2, the “anomaly” results.

To give some idea about the quality of approximations (7) and (8), we turn to those distributions whose exact values are known to us (Table 1). For the exponential at $n=998$, the approximation agrees with the exact value up to the 12th decimal place!

Table 1
Exact and approximate values of mse(M_n)

	Uniform distribution		Exponential distribution	
	$n=250$	$n=251$	$n=250$	$n=251$
Exact	0.000988111	0.000988142	0.00399609	0.00399590
Approximate	0.000988112	0.000988144	0.00399609	0.00399590

4. Examples

As was noted earlier, difference (9) depends only on f and the first two derivatives, providing us with an easy test of (3). For the special case of symmetric distributions, it is even simpler. There is no "anomaly" if $f''(0)$ is very negative ($< -4f^3(0)$). In other words, in order that an additional observation continues to contribute to the precision of the median it is necessary (but not sufficient) that f has a local mode at the origin. We shall see, nevertheless, that such is the case for most of the unimodal distributions, including the normal and the logistic. For the U-shaped densities ($f''(0) > 0$), it goes without saying that $\mathrm{mse}(M_n)$ always decreases with n in a zig-zag fashion.

The normal:

$$f(x) = (1/\sqrt{2\pi})\exp(-x^2/2).$$

The derivatives evaluated at $x = 0$ are

$$f = (2\pi)^{-1/2}, \quad f' = 0, \quad f'' = -(2\pi)^{-1/2}, \quad f''' = 0, \quad f'''' = 3(2\pi)^{-1/2}.$$

$$c_1 = \sqrt{2\pi}, \quad c_2 = 0, \quad c_3 = \tfrac{1}{6}(2\pi)^{3/2}, \quad c_4 = 0, \quad c_5 = \tfrac{7}{120}(2\pi)^{5/2}.$$

$$\mathrm{mse}(M_{2m}) = \tfrac{\pi}{4}m^{-1} + (-\tfrac{3}{8}\pi + \tfrac{1}{16}\pi^2)m^{-2} + (\tfrac{7}{16}\pi - \tfrac{3}{16}\pi^2 + \tfrac{13}{384}\pi^3)m^{-3} + \mathrm{o}(m^{-3}).$$

$$\mathrm{mse}(M_{2m+1}) = \tfrac{\pi}{4}m^{-1} + (-\tfrac{3}{8}\pi + \tfrac{1}{16}\pi^2)m^{-2} + (\tfrac{9}{16}\pi - \tfrac{1}{4}\pi^2 + \tfrac{13}{384}\pi^3)m^{-3} + \mathrm{o}(m^{-3}).$$

It follows that $\mathrm{mse}(M_{2m}) - \mathrm{mse}(M_{2m+1}) = \frac{\pi}{16}(\pi - 2)m^{-3} + \mathrm{o}(m^{-3})$.

Since the coefficient of the m^{-3} term is positive, each additional observation from the normal distribution continues to improve the precision. The common wisdom "it does not pay to base the median on an odd number of observations" does not even hold for the normal case.

The exponential:

$$f(x) = \exp(-x), \; x > 0.$$

$$f = f'' = f'''' = \tfrac{1}{2}, \quad f' = f''' = -f = -\tfrac{1}{2}.$$

$$c_1 = 2, \quad c_2 = 2, \quad c_3 = \tfrac{8}{3}, \quad c_4 = 4, \quad c_5 = \tfrac{32}{5}.$$

$$\mathrm{mse}(M_{2m}) = \tfrac{1}{2}m^{-1} - \tfrac{1}{16}m^{-2} + \tfrac{17}{96}m^{-3} + \mathrm{o}(m^{-3}),$$

$$\mathrm{mse}(M_{2m+1}) = \tfrac{1}{2}m^{-1} - \tfrac{1}{16}m^{-2} - \tfrac{19}{96}m^{-3} + \mathrm{o}(m^{-3}),$$

Thus $\mathrm{mse}(M_{2m}) - \mathrm{mse}(M_{2m+1}) = \frac{3}{8}m^{-3} + \mathrm{o}(m^{-3})$.

Again, M_{2m} is no better than M_{2m+1}. In fact, from Section 2 we know this as a fact not just asymptotically, but for all finite $m \geqslant 1$.

The logistic:

$$f(x) = \mathrm{e}^{-x}(1 + \mathrm{e}^{-x})^{-2}.$$

$$f = \tfrac{1}{4}, \quad f' = 0, \quad f'' = -\tfrac{1}{8}, \quad f''' = 0, \quad f'''' = \tfrac{1}{4}.$$

$$c_1 = 4, \quad c_2 = 0, \quad c_3 = \tfrac{16}{3}, \quad c_4 = 0, \quad c_5 = \tfrac{64}{5}.$$

$$\mathrm{mse}(M_{2m}) = 2m^{-1} - m^{-2} + \tfrac{4}{3}m^{-3} + \mathrm{o}(m^{-3}),$$

$$\mathrm{mse}(M_{2m+1}) = 2m^{-1} - m^{-2} + \tfrac{1}{3}m^{-3} + \mathrm{o}(m^{-3}),$$

$$\mathrm{mse}(M_{2m}) - \mathrm{mse}(M_{2m+1}) = m^{-3} + \mathrm{o}(m^{-3}).$$

The variance decreases monotonely with n, same as the previous examples.

The uniform:

$$f(x) = 1,\ \ 0 \leqslant x \leqslant 0.$$

$$f = 1, \quad f' = f'' = f''' = f'''' = 0.$$

$$c_1 = 1, \quad c_2 = c_3 = c_4 = c_5 = 0.$$

$$\mathrm{mse}(M_{2m}) = \tfrac{1}{8}m^{-1} - \tfrac{3}{16}m^{-2} + \tfrac{7}{32}m^{-3} + \mathrm{o}(m^{-3}),$$

$$\mathrm{mse}(M_{2m+1}) = \tfrac{1}{8}m^{-1} - \tfrac{3}{16}m^{-2} + \tfrac{9}{32}m^{-3} + \mathrm{o}(m^{-3}),$$

$$\mathrm{mse}(M_{2m}) - \mathrm{mse}(M_{2m+1}) = -\tfrac{1}{16}m^{-3} + \mathrm{o}(m^{-3}).$$

As expected, the leading coefficient is negative. This is an example where the average of two central values is more stable than a single central value. It does not pay to base the median on an odd number of observations!

5. Small sample results

Since our results are asymptotic, it is important to know how good they are for small samples. Exact values of expectation and covariance of the order statistics are available for the normal for samples up to $n \leqslant 20$, (see, for instance, Beyer, 1968). For the theoretical values of the exponential order statistics, see Section 2. In Table 2 we show the exact values of mse against the approximations of Theorem 1 for the normal and the exponential distribution.

Table 2
Exact and approximate values of $\mathrm{mse}(M_n)$, small n

	Normal distribution		Exponential distribution	
n	Exact value	Approximate value	Exact value	Approximate value
5	0.28683366	0.29606689	0.22174465	0.20963542
6	0.21474267	0.22068260	0.16618910	0.16628086
7	0.21044686	0.21238070	0.15509180	0.15239198
8	0.16818086	0.17023391	0.12384180	0.12386068
9	0.16610128	0.16673155	0.11891158	0.11800130
10	0.13832644	0.13921846	0.09891158	0.09891667
11	0.13716243	0.13742525	0.09630437	0.09591667
12	0.11751619	0.11796500	0.08241548	0.08241705
13	0.11679900	0.11692726	0.08087292	0.08068094
14	0.10216784	0.10241799	0.07066883	0.07066934
15	0.10169465	0.10176448	0.06968166	0.06957604
16	0.09037521	0.09052558	0.06186916	0.06186930
17	0.09004658	0.09008778	0.06119969	0.06113688
18	0.08102851	0.08112430	0.05502685	0.05502686
19	0.08079098	0.08081682	0.05455211	0.05451246
20	0.07343703	0.07350093	0.04955211	0.04955208

The agreement is quite good for practical purposes. Notice also that the mse actually decreases monotonely in n for the normal distribution, regardless of parity, at least up to $n = 20$. We suspect this to be true for larger n. Simulation results seem to confirm this.

Remark 1. To the best of our knowledge mse(M_n) has not been calculated to the third-order term before. It is straightforward, but extremely laborious. In view of the importance of sample medians, we felt that the labor is worthwhile. Besides its role in breaking the longheld misconception the approximation, which is rather accurate, should prove useful in its own right in the study of robust statistics.

Remark 2. Should the coefficient of the m^{-3} term in (9) turn out to be zero, a higher-order expansion (God forbids!) will of course be called for. We have found just such an interesting example.

Sine distribution:

$$f(x) = \tfrac{1}{2}\cos x, \ -\tfrac{\pi}{2} \leqslant x \leqslant \tfrac{\pi}{2}.$$

$$f = \tfrac{1}{2} = -f'' = f'''', \quad f' = f''' = 0.$$

$$c_1 = 2, \quad c_2 = 0, \quad c_3 = \tfrac{4}{3}, \quad c_4 = 0, \quad c_5 = \tfrac{12}{5}.$$

$$\mathrm{mse}(M_{2m}) = \mathrm{mse}(M_{2m+1}) = \tfrac{1}{2}m^{-1} - \tfrac{1}{2}m^{-2} + \tfrac{11}{24}m^{-3} + \mathrm{o}(m^{-3}).$$

Some properties of this distribution were reported by Burrows (1986). We have simply added a new one.

References

Beyer, W.R., 1968. Handbook of Tables for Probability and Statistics, 2nd ed. Chem. Rubber Co., Cleveland, OH.

Burrows, P.M., 1986. Extreme statistics from the sine distribution. Amer. Statist. 40, 216–217.

Cabrera, J., Maguluri, G., Singh, K., 1994. An odd property of the sample median, Statist. Probab. Lett. 19, 349–354.

David, H.A., 1981. Order Statistics, 2nd ed., Wiley, New York.

Hodges, J.L. Jr., Lehmann, E.L., 1967. On medians and quasi medians. J. Amer. Statist. Assoc. 62, 926–931.

Mudholkar, G.S., Hutson, A.D., 1997. Asymmetric quasimedians: remarks on an anomaly. Statist. Probab. Lett. 32, 261–268.

Oosterhoff, J., 1994. Trimmed mean or sample median?. Statist. Probab. Lett. 20, 401–409.

Stigler, S.M., 1977. Fractional order statistics, with applications. J. Amer. Statist. Assoc. 72, 544–550.