Option-3 Median Measurement Scheme in Outlier Reduction

M. C. K. Yang[1,3], Y. Y. Namgung[5], T. Kang[1,3], I. Magnusson[3,4] and A. Hefti[2,3]

[1]Department of Statistics, [2] Department of Periodontology, [3] Periodontal Disease Research Center, [4]Department of Oral Biology, University of Florida, Gainesville, Florida 32610,

[5] National Statistical Office, Seoul, 135-723, Republic of Korea,

SUMMARY

When measurements are subject to rare but large errors, it is better to measure twice instead of once, and if the two measurements differ too much, take a third measurement. This is called the option-3 scheme. This paper shows that in this scheme, using the median of the three measurements is better than using the average of the two closest. To reach the maximum sample size benefit, the threshold for taking the third measurement is approximately 3 times the measurement error standard deviation.

1. INTRODUCTION

When measuring an unknown quantity, a single measurement may not be adequate. Two measurements are much better, but what if the two differ too much? Obviously, in that case a third measurement should be taken. But how much is too much and how should the three measurements be combined? These questions are often asked when probing of periodontal attachment level is performed to examine periodontal disease. Periodontal disease of some form is present in virtually all individuals worldwide and continues to be one of mankind's major dental problems. According to a 1987 NIH report, 95% of adults age 65 or over in the United States have experienced some loss of periodontal attachment around their teeth. Using a probe to measure the attachment level is one of the most common practices in assessing the severity of the attachment loss.[1-3] Due to its importance, there is a vast literature on the accuracy of the attachment measurements.[3-7]

However, at the current state of the art, even with the best instrument and under very carefully arranged clinical conditions, outliers cannot be totally eliminated.[4,5,7] A measurement outlier can cause a false alarm in claiming attachment breakdown and subsequent unnecessary treatment and cost. A remedial scheme has been proposed in the literature. Two probings are taken initially from a site and if the two probing values are not close enough, a third probing is taken, and the average of the closest two is used as the estimate of the "true" attachment level. However, there is no consensus on the definition of closeness. Osborn et al.[8] used 1mm as the threshold (symbol δ will be used), Best et al.[9] chose δ =3mm, and Marks et al.[10] suggested δ =1.1mm. Namgung and Yang[11] published a theoretical study on this method and named it the option-3 scheme, because the third measurement is optional. In contrast, there is the "take-3" method which always takes 3 measurements and uses the closest two or the median to estimate the true value.[12–14] It was found that the option-3 method, as suggested in the dental literature, was satisfactory in eliminating outliers. Although the optimal choice of δ depends on the magnitude and occurrence frequency of the outliers, δ is not very sensitive to the two quantities. Letting $\delta = 3.5$ times the measurement standard deviation is a good choice. One reason that the option-3 method is preferred to the take-3 method is the time and cost involved in taking the third measurement. Also, since pain can intensify with repeated probings at the same site, a dentist wants to minimize the number of probings. Because outliers in measurements can occur quite often[15], take-3 may be replaced by the option-3 scheme in many other medical measurements.

The purpose of this paper is to investigate the median option of the option-3 methods. It is found that under very mild conditions the median choice is better than the closest-2 choice in estimating the true value of the measured quantity.

2. ANALYTIC FORMULATION

Let the true value of a quantity to be measured be μ. Two measurements, x_1 and x_2, are taken and a threshold δ is predetermined. If $|x_1 - x_2| < \delta$, then the estimate for μ is $\hat{\mu} = (x_1 + x_2)/2$;

else, a third measurement x_3 is taken and $\hat{\mu}$ =Median (x_1, x_2, x_3). We call this scheme the option-3 median method and model it by letting

$$x_i = \mu + \epsilon_i \qquad i = 1, 2, 3$$

where the ϵ_i's are independent and identical measurement errors with 0 mean. Following the general outlier modeling,[16] we let the distribution of the ϵ_i's be

$$\epsilon = pF + qG;\ 0 \leq p \leq 1,\ q = 1 - p, \tag{1}$$

where F is the basic error, G is the contaminant, and q is the occurrence frequency of G. Let $N(\mu, \sigma^2)$ denote the normal distribution with mean μ and variance σ^2. We assume $F \sim N(0, \sigma^2)$ with σ being known. This assumption agrees with a usual measurement model and is perfect in the attachment probing case.[5,7] The distribution of G may not be easy to identify due to data sparsity, but it is usually assumed to be a symmetric distribution centered at 0.[16] First, we let G be normal with a much larger variance, i.e., $G \sim N(0, k^2\sigma^2)$, where $k >> 1$. Without loss of generality, we may let $\sigma = 1$. Thus, the density function of ϵ becomes

$$f(\epsilon) = \frac{p}{\sqrt{2\pi}} e^{\frac{-\epsilon^2}{2}} + \frac{q}{\sqrt{2\pi}k} e^{\frac{-\epsilon^2}{2k^2}}\ . \tag{2}$$

The option-3 median estimate of μ, denoted by $\hat{\mu}$, is given by

$$\hat{\mu} = \begin{cases} \frac{x_1 + x_2}{2} & \text{if } |x_1 - x_2| < \delta \\ \text{Median}(x_1, x_2, x_3) & \text{if } |x_1 - x_2| \geq \delta\ . \end{cases} \tag{3}$$

Equation (**??**) depends on the threshold value δ, which will be chosen so that it minimizes the mean square error (MSE) $E(\hat{\mu} - \mu)^2$ (same as the variance of $\hat{\mu}$, since $E(\hat{\mu}) = \mu$ can be proven by symmetry). Obviously,

$$E(\hat{\mu} - \mu)^2 = E\left[(\hat{\mu} - \mu)^2 I_{|x_1 - x_2| < \delta}\right] + E\left[(\hat{\mu} - \mu)^2 I_{|x_1 - x_2| \geq \delta}\right]\ , \tag{4}$$

where I_A is the indicator function which is 1 if A is satisfied and 0 otherwise.

The first part in (**??**) can be represented as

$$E\left[(\hat{\mu}-\mu)^2 I_{|x_1-x_2|<\delta}\right] = \iint_{|x_1-x_2|<\delta}\left(\frac{x_1+x_2}{2}-\mu\right)^2 f(x_1)f(x_2)dx_1\,dx_2$$
$$= \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\sqrt{a_i a_j}}{4\sqrt{\pi}h_1^{5/2}}\left\{2h_1\Psi_1(-\delta,\delta,h_3,0)+h_2^2\,\Psi_3(-\delta,\delta,h_3,0)\right\}. \quad (5)$$

where $a_1 = 1/2,\ a_2 = 1/(2k^2),\ h_1 = a_i + a_j,\ h_2 = a_i - a_j,\ h_3 = a_i a_j/h_1$ and Ψ's are defined in the Appendix where derivations are briefly described. The second part of (**??**) can be evaluated as follows.

$$E\left[(\hat{\mu}-\mu)^2 I_{|x_1-x_2|\geq\delta}\right]$$
$$= \iiint_{|x_1-x_2|\geq\delta}(\mathrm{Median}(x_1,x_2,x_3)-\mu)^2 f(x_1)f(x_2)f(x_3)dx_1\,dx_2\,dx_3$$
$$= 2\sum_{i,j,k}^{2}\frac{\sqrt{a_i a_j a_k}}{\pi\beta^{5/2}}\int_{\delta}^{\infty} g_1(t)\ e^{\frac{-a_i a_j t^2}{(\beta-a_k)}}dt$$
$$+ \sum_{i,j,k}^{2}\frac{\sqrt{a_i a_j a_k}}{\pi\beta^{5/2}}\left(\int_0^{\delta} g_2(t)\ e^{-\frac{a_i a_k}{(\beta-a_j)}t^2}dt + \int_0^{\delta} g_3(t) e^{-\frac{a_i a_k}{(\beta-a_j)}t^2}dt\right),$$
(6)

where $\beta = a_i + a_j + a_k$ and the $g_i(t)$'s are defined in the Appendix. Equation (**??**) now involves only single integrals and can be computed with great accuracy.

3. BENEFITS IN USING THE OPTION-3 MEDIAN SCHEME

A comparison of the closest-2 and median choices of option-3 for various q's and δ's is given in Table 1. The values of the former were computed based on the formulae described in Namgung and Yang[11]. We can see from this table that the median choice is better than the closest-2 choice in almost all the cases. Thus, for threshold δ in the optimal range from 3 to 5 times the measurement standard deviation, the closest-2 estimate should be replaced by the median estimate.

It is relevant to investigate what is the optimal δ when the median is used. In Table 2, one part of our simulation study is presented. The data in Table 2 also confirm the analytically computed values in Table 1.

Table 2 reveals another interesting difference between the closest-2 and the median choices. While the option-3 closest-2 MSE reaches minimum at δ at approximately 3.5, the MSE of the option-3 median is a monotonic increasing function of δ. Because the median is not sensitive to the *numerical value* of an outlier, it turns out that the third measurement always helps. However, this does not mean that when the median choice is used, we should set $\delta = 0$, i.e., use the take-3 scheme. The gain in accuracy may not be worth the cost. In order the see the trade-oft, we used the sample size benefit as a measure.[11] The sample size benefit is defined as

$$\text{sample size benefit } = \frac{\sigma_0^2}{\sigma_{\hat{\mu}}^2} - n_{\hat{\mu}}, \tag{7}$$

where σ_0^2 is the variance when the target quantity is measured once, $\sigma_{\hat{\mu}}^2$ is the variance when a sampling scheme is used and $n_{\hat{\mu}}$ is the average sample size. Since we expect the variance of $\hat{\mu}$ to be $\sigma_0^2/n_{\hat{\mu}}$ when the sample mean is used, the value in (**??**) is the gain (or loss if negative) in sample size. Table 3 show a typical relation between δ and the sample size benefit. More detailed computation shows that the δ for the maximum sample size benefit is around 3. A summary of the optimal δ's is given in Table 4.

In Table 4, it is a little surprising to see that the optimal δ for both estimates are very similar. Thus, the previously suggested δ=3.5 is still a reasonable choice, although $\delta = 3$ is better for the median estimate. It is to be expected that the difference in optimum δ will have only minor practical implication.

As mentioned before, the outlier distribution G in (**??**) is not easy to identify. Table 5 shows a comparison when G is a normal distribution, a shorter tail uniform distribution, and a longer tail double exponential distribution (D-exp). The variances of G were kept the same for all three distributions. From this table and from intuition, it is apparent that the benefit is higher for a long tail than for a short tail distribution. However, the difference seems not great enough to cause concerns in practice.

4. CONCLUDING REMARKS

Our results show that option-3 median offers some advantage over the closest-2 choice. To reach the maximum sample size benefit, the threshold δ for initiating a third measurement is approximately 3 times the measurement standard deviation.

We designed a clinical experiment to test this result. Twelve subjects (2 females and 10 males; ages ranging from 26 to 34 years) with excellent dental health from a population of dental and graduate students volunteered to participate in a study involving four visits over a period of 15 weeks. Healthy subjects were chosen to ensure that there would be little change in the attachment level during the study period. Two types of probes and two examiners were used to investigate instrument-induced and inter-examiner variation. A probing scheme was considered better when it produced more consistent results at different visits. Thus, if there is inconsistency in attachment level among visits from a periodontal patient in a real clinical setting, then the difference is likely due to real change rather than measurement error. One measurement, two measurements, option-3 closest-2 and option-3 median schemes were compared in a total of 1,728 sites (6 sites per tooth). We found that the option-3 median was indeed more consistent than the option-3 closest-2 choice, followed by the average of two measurements method and one measurement per site. Details of this study can be found in Johnson et al.[17]

Closest-2 and median are only two of many choices. Using the average of the three measurements(average-3) is another alternative. We have done a simulation study on the linear combination of the order statistics of the three, i.e, let the three measurements be $x_{(1)} \leq x_{(2)} \leq x_{(3)}$ and let the estimate be

$$\hat{\mu} = \sum_{i=1}^{3} w_i x_{(i)}, \tag{8}$$

where w_i are weights with $w_i \geq 0$ and $w_1 + w_2 + w_3 = 1$ (necessary for unbiasness). When $w_2 = 1$, it is the median estimate and when $w_1 = w_2 = w_3 = 1/3$ it is the average of the three. By symmetry, we can let $w_1 = w_3$. Thus, the weights in (**??**) contain one variable. A simulation study showed that the median is indeed the best choice for reasonable choices of δ (in [3, 5]), k (in [3, 5]) and q (in [0.02, 0.1]). As expected, when k is small (≤ 3) and q very small (≤ 0.02), average-3 is better than the median, but the difference is insignificant as compared to large q. For example, when δ=3 and $k = 3$, the MSE of the option-3 median to that of the option-3 average-3 is $1 : 0.98$ at $q = 0.01$, but this trend is reversed at $q = 0.05$ when the ratio is $1 : 1.01$. Again, when δ=3 but let $k = 5$, the ratios are $1 : 1.11$ and $1 : 1.42$ for $q = 0.01$ and 0.1 respectively. We can conclude that using the median is the best in (**??**) for large contaminant error, but is reasonably good even for small contaminants. Unfortunately, (**??**) does not cover the closest-2 type of estimates. It is a theoretical challenge on what is the best way to use the three measures. This problem does not seem to have been investigated for the take-3 method either.

APPENDIX: DERIVATION OF (??) and (??)

The derivation of (**??**) is,

$$\iint_{|x_1-x_2|<\delta}\left(\frac{x_1+x_2}{2}-\mu\right)^2 f(x_1)f(x_2)dx_1\,dx_2$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{2}\iint_{|z_1-z_2|<\delta}\left(\frac{z_1+z_2}{2}\right)^2\frac{\sqrt{a_ia_j}}{\pi}e^{-(a_iz_1^2+a_jz_2^2)}dz_1dz_2. \tag{9}$$

Using the variable transformation $t = z_1 + z_2$ and $u = z_1 - z_2$; $|u| < \delta$ and $-\infty < t < \infty$ and integrate out t, we have (**??**). The Ψ functions are defined as:

$$\begin{aligned}\Psi_1(a,b,c,d) &= \int_a^b e^{-c(x-d)^2}dx \\ &= \sqrt{\frac{\pi}{c}}\left[\Phi(\sqrt{2c}(b-d)) - \Phi(\sqrt{2c}(a-d)\right],\end{aligned}$$

$$\Psi_2(a,b,c,d) = \int_a^b xe^{-c(x-d)^2}dx$$

$$= \frac{1}{2c}\left[e^{-c(a-d)^2-c(b-d)^2}\right] + d\Psi_1(a,b,c,d)\,,$$

$$\begin{aligned}
\Psi_3(a,b,c,d) &= \int_a^b x^2 e^{-c(x-d)^2}dx \\
&= \frac{1}{2c}\left[(a+d)e^{-c(a-d)^2} - (b+d)e^{-c(b-d)^2}\right] + \\
&\quad \frac{1}{2c}\Psi_1(a-d,b-d,c,0) + d^2\Psi_1(a,b,c,d)\,, \qquad (10)
\end{aligned}$$

where Φ is the cumulative standard normal distribution function.

The derivation of (**??**) is based on,

$$\iiint_{|x_1-x_2|\geq\delta} (\text{Median}(x_1,x_2,x_3)-\mu)^2 f(x_1)f(x_2)f(x_3)dx_1\,dx_2\,dx_3$$

$$\begin{aligned}
= \; & 2\iiint_{\substack{|z_1-z_2|\geq\delta \\ z_1=\text{Median}(z_1,z_2,z_3)}} z_1^2 f(z_1)f(z_2)f(z_3)dz_1\,dz_2\,dz_3 \\
& + \iiint_{\substack{|z_1-z_2|\geq\delta \\ z_3=\text{Median}(z_1,z_2,z_3)}} z_3^2 f(z_1)f(z_2)f(z_3)dz_1\,dz_2\,dz_3
\end{aligned} \qquad (11)$$

where $f(z)$ is defined in (**??**). The range of the first integral, $|z_1-z_2|\geq\delta$ and $z_1 = \text{Median}(z_1,z_2,z_3)$, can be presented by the following two parts:

(i) $|z_1-z_2|\geq\delta,\;\; z_2<z_1<z_3$, with variable transformation

$t=z_1-z_2,\;\; u=z_1,\;\; v=z_3-z_1 \quad \Longrightarrow\;\; t\geq\delta,\;\; -\infty<u<\infty,\;\; v>0,$

and (ii) $|z_1-z_2|\geq\delta,\;\; z_3<z_1<z_2$, with variable transformation

$t=z_1-z_3,\;\; u=z_1,\;\; v=z_3-z_1 \quad \Longrightarrow t>0,\;\; v>\delta,\;\; -\infty<u<\infty.$

The first integral in (**??**) can then be shown to be

$$2\sum_{i,j,k}^{2}\int_\delta^\infty\int_0^\infty\int_{-\infty}^\infty \frac{\sqrt{a_ia_ja_k}}{\pi^{3/2}}u^2 e^{-\beta(u-\frac{a_jt-a_kv}{\beta})^2}due^{-\frac{a_k(\beta-a_k)}{\beta}(v+\frac{a_jt}{\beta-a_k})^2-\frac{a_ia_j}{(\beta-a_k)}t^2}dvdt. \qquad (12)$$

The two integrals with ranges from $-\infty$ to ∞ and from 0 to ∞ can be simplified into standard normal functions. The range for the second integral in (**??**) is divided as the following two parts:

(i) $|z_1 - z_2| \geq \delta,\ z_1 < z_3 < z_2$, with variable transformation

$$t = z_3 - z_1,\ u = z_3,\ v = z_2 - z_3$$
$$\Longrightarrow \begin{cases} -\infty < u < \infty \quad v \geq \delta - t \quad 0 < t < \delta \\ \\ -\infty < u < \infty \quad v > 0 \qquad\quad t > \delta, \end{cases}$$

and (ii) $|z_1 - z_2| \geq \delta,\ z_2 < z_3 < z_1$. This is the same as (i), when z_1 and z_2 are switched.

Other derivations are similar to that of (**??**). The three g functions in (**??**) are:

$$\begin{aligned}
g_1(t) &= (\beta + 2a_j^2t^2)\Psi_1(0,\infty,\frac{a_k(\beta-a_k)}{\beta},-\frac{a_jt}{(\beta-a_k)}) \\
&\quad -4a_ja_kt\Psi_2(0,\infty,\frac{a_k(\beta-a_k)}{\beta},-\frac{a_jt}{(\beta-a_k)}) + 2a_k^2\Psi_3(0,\infty,\frac{a_k(\beta-a_k)}{\beta},-\frac{a_jt}{(\beta-a_k)}) \\
g_2(t) &= (\beta + 2a_i^2t^2)\Psi_1(\delta-t,\infty,\frac{a_j(\beta-a_j)}{\beta},-\frac{a_it}{\beta-a_j}) \\
&\quad -4a_ia_jt\Psi_2(\delta-t,\infty,\frac{a_j(\beta-a_j)}{\beta},-\frac{a_it}{\beta-a_j}) + 2a_j^2\Psi_3(\delta-t,\infty,\frac{a_j(\beta-a_j)}{\beta},-\frac{a_it}{\beta-a_j}) \\
g_3(t) &= (\beta + 2a_i^2t^2)\Psi_1(0,\infty,\frac{a_j(\beta-a_j)}{\beta},-\frac{a_it}{\beta-a_j}) \\
&\quad -4a_ia_jt\Psi_2(0,\infty,\frac{a_j(\beta-a_j)}{\beta},-\frac{a_it}{\beta-a_j}) + 2a_j^2\Psi_3(0,\infty,\frac{a_j(\beta-a_j)}{\beta},-\frac{a_it}{\beta-a_j}),
\end{aligned}$$

where the Ψ functions are defined in (**??**).

REFERENCES

1. Imrey P. B. 'Considerations in the statistical analysis of clinical trials in periodontitis', *Journal of Clinical Periodontology*, **13,** 517-528 (1986).

2. Birkedal-Hansen, H., Lantz, M. S. and Cogen, R. B. 'A critical assessment of the burst hypothesis', *Periodontology Today*, Int. Congr., Zurich, 227-235 (1988).

3. Janssen P. T. M., Faber J. A. J., van Palenstein Helderman W. H. 'Accuracy of repeated single versus averages of repeated duplicates of probing depth measurements', *Journal of Clinical Periodontology*, **15**, 569-574 (1988).

4. Goodson J. M. 'Clinical measurements of periodontitis', *Journal of Clinical Periodontology*, **13**, 446-455 (1986).

5. Gunsolley J. C., Best A. M. 'Change in attachment levels', *Journal of Periodontology* **59**, 450-456 (1988).

6. Karim M., Birek P., McCullock C. A. 'Controlled force measurements of gingival attachment levels made with the Toronto automated probe using electronic guidance', *Journal of Clinical Periodontology*, **17**, 594-600 (1990).

7. Yang M. C. K., Marks R. G., Magnusson I., et al. 'Reproducibility of an electronic probe in relative attachment level measurements', *Journal of Clinical Periodontology*, **19**, 541-548 (1992).

8. Osborn J, Stoltenberg J. H. B., Aeppli D. et al. 'Comparison of measurement variability using a standard and constant force periodontal probe', *Journal of Periodontology*, **61**, 496-503 (1990).

9. Best A. M., Burmeister J. A., Gunsolley J. C., et al. 'Reliability of attachment loss measurements in a longitudinal clinical trial'. *Journal of Clinical Periodontology*, **17**, 564-569 (1990).

10. Marks R. G., Low S. B., Taylor M., et al. 'Reproducibility of attachment level measurements with two models of the Florida Probe', *Journal of Clinical Periodontology*, **18**, 780-784 (1990).

11. Namgung Y. Y. and Yang M. C. K. 'Outlier reduction by an option-3 measurement scheme', *Biometrics*, **50**, 173-182 (1994).

12. Seth, G. R. 'On the distribution of the two closest among a set of three observations', Annals of Mathematical Statistics', **21**, 298-301 (1950).

13. Leiblein, J. 'Properties of certain statistics involving the closest pair in a sample of three observations', *Journal of Research of the National Bureau of Standards*, **48**, 255-268 (1952).

14. Stefanski, L. A. and Meredith M. 'Robust estimation of location in samples of size three', *Communication of Statistics - Theory and methods*, **15(10)**, 2921-2933 (1986).

15. Chatfield, C. 'Avoiding statistical pitfalls (with discussion)', *Statistical Science* **6**, 240-268

(1991).

16. Beckman R. J. and Cook, R. D. 'Outliers', *Technometrics*, **25**, 119-163 (1983).

17. Johnson, P., Yang, M. C. K., Magnusson, I., et al., 'Evaluating the merits of option-3 methods in relative attachment level measurement in a healthy population,' *Journal of Periodontal Research*, **32**, 256-263 (1997).

Table 1. Comparison of MSE by closest-2 and median choices in option-3 schemes

$\epsilon \sim pN(0,1) + qN(0,k^2)$; δ= threshold to initiate a third measurement

q	δ	$k=3$		$k=4$		$k=5$	
		Closest-2	Median	Closest-2	Median	Closest-2	Median
0.01	3	0.546	0.524	0.547	0.527	0.546	0.528
	4	0.526	0.521	0.528	0.525	0.527	0.525
	5	0.526	0.525	0.530	0.531	0.530	0.532
0.025	3	0.569	0.547	0.572	0.554	0.573	0.559
	4	0.556	0.551	0.562	0.560	0.563	0.565
	5	0.564	0.563	0.576	0.578	0.580	0.584
0.05	3	0.612	0.588	0.624	0.605	0.635	0.619
	4	0.610	0.603	0.628	0.624	0.639	0.638
	5	0.640	0.628	0.660	0.660	0.676	0.677
0.075	3	0.660	0.632	0.688	0.662	0.718	0.689
	4	0.668	0.658	0.704	0.694	0.734	0.721
	5	0.699	0.695	0.755	0.748	0.788	0.780
0.1	3	0.713	0.678	0.764	0.725	0.821	0.769
	4	0.729	0.714	0.790	0.770	0.848	0.814
	5	0.772	0.764	0.854	0.841	0.918	0.892

Table 2. The relation between MSE and δ, based on 10,000 simulations with $k = 4$. A : closest-2 estimate, B : median estimate

q	Est.	$\delta = 0.0$	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0
0.01	A	0.651	0.640	0.594	0.550	0. 528	0.529	0.539	0.549	0.557
	B	0.462	0. 502	0.529	0.528	0.524	0.529	0.540	0.549	0.557
0.05	A	0.711	0.702	0. 661	0. 630	0.631	0. 657	0.703	0.749	0.787
	B	0.524	0.562	0.591	0.602	0.622	0.655	0.702	0.747	0.786
0.1	A	0.834	0.826	0.791	0.774	0.797	0.850	0.933	0.1016	0.1086
	B	0.630	0.667	0.698	0.728	0.771	0.835	0.921	0.1006	0.1079

Table 3. Sample size benefit of the option-3 median scheme with $q = 0.05$. (A postive value stands for gain.)

δ	$k = 3$	$k = 4$	$k = 5$
1	-0.01	0.504	1.170
2	0.197	0.700	1.350
3	0.316	0.815	1.470
4	0.296	0.765	1.402
5	0.216	0.627	1.216
6	0.137	0.474	0.995
7	0.079	0.337	0.780
8	0.041	0.227	0.589
9	0.019	0.145	0.431

Table 4. Optimal δ and the corresponding variance

for the closest-2 and the median option-3 estimates.

k	q	closest-2		Median	
		Opt. δ	Variance	Opt. δ	Variance
3	0.01	4.5	0.524672	4.1	0.52159
	0.025	4.0	0.556232	3.7	0.54897
	0.05	3.6	0.606610	3.3	0.59099
	0.075	3.3	0.658318	3.1	0.63326
	0.1	3.1	0.713163	2.9	0.67529
4	0.01	4.3	0.526929	4.0	0.52455
	0.025	3.8	0.561301	3.5	0.55542
	0.05	3.4	0.620660	3.2	0.60728
	0.075	3.2	0.687690	2.9	0.66019
	0.1	3.0	0.764354	2.7	0.71745
5	0.01	4.3	0.526655	4.0	0.52586
	0.025	3.8	0.562736	3.5	0.55994
	0.05	3.4	0.631859	3.1	0.61963
	0.075	3.2	0.717237	2.9	0.68702
	0.1	3.03	0.820660	2.7	0.76200

Table 5. Sample size benefit of option-3 median scheme with normal and nonnormal contaminants (based on 5,000 simulations)

	Normal			Uniform			D-exp.		
q	k=3	k=4	k=5	k=3	k=4	k=5	k=3	k=4	k=5
0.010	0.063	0.181	0.344	−0.06	0.071	0.243	−0.01	0.107	0.266
0.025	0.165	0.443	0.820	0.015	0.317	0.701	0.130	0.402	0.773
0.050	0.321	0.817	1.471	0.137	0.664	1.305	0.335	0.856	1.517
0.075	0.454	1.114	1.958	0.244	0.917	1.678	0.533	1.218	2.099
0.100	0.563	1.340	2.302	0.327	1.072	1.866	0.692	1.508	2.535