MathSoft

Research Report
No. 85

Computation of Weighted Functional Statistics Using Software That Does Not Support Weights

Stephen J. Ellis and Tim C. Hesterberg

Revision 1, Revision Date: October 12, 1999

Acknowledgments: This work was supported by NSF Phase I SBIR Award No. DMI-9861360.

MathSoft, Inc.
1700 Westlake Ave. N, Suite 500
Seattle, WA 98109–9891, USA
Tel: (206) 283–8802
FAX: (206) 283–6310

E-mail: sellis@statsci.com
timh@statsci.com
Web: www.statsci.com/Hesterberg/tilting

Computation of Weighted Functional Statistics Using Software That Does Not Support Weights

S. J. Ellis and T. C. Hesterberg

Abstract

We discuss methods for calculating statistics for weighted samples using software that does not support weights. Such samples arise in survey sampling with unequal probabilities, importance sampling, and bootstrap tilting. The software might not support weights for reasons of efficiency, simplicity, or because it was quicker to write the software without supporting weights. We discuss several techniques, both deterministic and random, for efficiently approximating the answer that would be obtained if the software supported weights.

Key Words: weights, function evaluation.

1 Introduction

In this report we propose methods to evaluate statistics for data that are associated with unequal probabilities (weights), using statistical software functions that are not written to allow weights. Data with weights arise in a number of contexts, including survey sampling with unequal probabilities, importance sampling (Hesterberg (1995b)), and bootstrap tilting inferences (Efron (1981)). We assume that the software could have been written to handle weights — that there is no intrinsic reason that the statistic being calculated could not be calculated with weights — but was not, possibly for reasons of efficiency, simplicity, or to reduce the development effort.

We assume that the statistic θ to be calculated is functional (Efron and Tibshirani (1993)), i.e. that $\theta(x_1, \ldots, x_n)$ depends on that data only via the corresponding empirical distribution that places mass $1/n$ on each of the observed data values x_i, $i = 1, \ldots, n$ (which may be multivariate). For example, the usual sample variance $s^2 = (n-1)^{-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$ is not a functional statistic, whereas $n^{-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$ is functional — it is the variance of the distribution with mass $1/n$ at each observed data point.

Note that a functional statistic gives the same result if applied to the original data or if applied to the dataset created by repeating each original observation the same number of times, e.g.

$$\theta(x_1, x_2, \ldots, x_n) = \theta(x_1, x_1, x_2, x_2, \ldots, x_n, x_n)$$

because both datasets (original and created) correspond to the same empirical distribution.

Let $\vec{x} = (x_1, \ldots, x_n)$ denote the original data, and $\vec{w} = (w_1, \ldots, w_n)$ a corresponding vector of weights, which are non-negative and sum to 1. Let $\theta(\vec{x}, \vec{w})$ denote the statistic calculated for the distribution with mass w_i on observation x_i.

Our goal is to approximate θ using a software function θ^s that does not handle weights. Our basic method is to repeat each original observation x_i a number of times M_i, with M_i approximately proportional to w_i for $i = 1, \ldots, n$,

$$\theta(\vec{x}, \vec{w}) \doteq \theta^s(\underbrace{x_1, \ldots, x_1}_{M_1 \text{times}}, \underbrace{x_2, \ldots, x_2}_{M_2 \text{times}}, \ldots, \underbrace{x_n, \ldots, x_n}_{M_n \text{times}})$$

If θ is a continuous function of $\vec{w}$, then the approximation can be made arbitrarily accurate by choosing sufficiently large M and letting $M_i = \lfloor M w_i \rceil$, where $\lfloor c \rceil$ denotes c rounded to the nearest integer.

However, this may be slow if $M^* = \sum M_i$ is large and calculating the statistic θ^s requires computational time which is super-linear in the number of observations. We therefore also propose a number of methods based on averaging a number of evaluations of the statistic, each with relatively small M^*, and with repetition values M_i chosen randomly with expected proportions $E(M_i/M^*) \doteq w_i$.

We propose a number of methods in Section 2, and compare their performance in a simulation study in Section 3.

2 Description of Methods

We describe five methods in this section. The first is deterministic. The others are random, based on averaging results from K random sets of M_i, for some $K \geq 1$.

Method 1 Our first method is essentially the "basic" method described above: For some M with $M \geq n$, let $M_i = \lfloor M w_i \rceil$, and let the estimate of $\theta(\vec{x}, \vec{w})$ be $\theta^s(\vec{x}, \vec{M})$, i.e. θ^s applied to the the sample created by repeating observation x_i M_i times.

Note that $|w_i - M_i/M^*| \leq |w_i - M_i/M| + |M_i/M - M_i/M^*| \leq 1/(2M) + n/(2M)$, so that the approximation is consistent as $M \to \infty$ (for fixed n and $\vec{x}$) if θ is a continuous function of $\vec{w}$.

Optionally, the values M_i may be adjusted so the total sample size is M. We did not do this.

A variation on this method is to choose the value of M within a pre-specified range (e.g. $M \leq M_{\max}$), in order to minimize the sum of rounding errors $\sum |w_i - M_i/M^*|$. The range may be chosen based on how much computing time is available.

Method 2 For some M, create a sample of size M by sampling the original data with replacement with probabilities $\vec{w}$. Then $\vec{M}$ has a multinomial distribution.

If $K > 1$, generate K such samples, independently. The final estimate is

$$\hat{\theta}(\vec{x}, \vec{w}) = K^{-1} \sum_{k=1}^{K} \theta^s(\vec{x}, \vec{M}_k) \tag{1}$$

Method 3 This method is similar to Method 2, except that the variability in the random sampling is reduced. Decompose Mw_i into integer and fractional parts $M_i^I = \lfloor Mw_i \rfloor$ and $M_i^F = Mw_i - M_i^I$. First, x_i is deterministically included in each created sample M_i^I times. Then the remaining $M - \sum M_i^I = \sum M_i^F$ observations are allocated by sampling with replacement with probabilities M_i^F. By sampling only the fractional parts, the overall variability is substantially reduced.

The random part of the procedure is repeated K times, and the final estimate has the same form as (1).

Method 4 Here the overall variability is reduced further by generating all K replicates "at once" (not independently). The integer and fractional parts M_i^I and M_i^F are computed as in Method 3, and x_i is again deterministically included in each created sample M_i^I times. Then decompose KM_i^F into integer and fractional parts, and generate a temporary sample of size $K \sum M_i^F$ using a combination of deterministic allocation and random sampling. This temporary sample is randomly reordered and split into K parts of with equal lengths $\sum M_i^F$. Those parts are appended to the deterministic observations to create the K final created samples.

The final estimate has the same form as (1). The K values $\theta^s(\vec{x}, \vec{M}_k)$ are not independent, but rather are negatively correlated (in most applications).

Method 5 This method reduces the variability in the number of repetitions allocated to *similar* observations within the same sample, using the empirical influence function to quantify similarity. Let

$$L_i = \lim_{\epsilon \to 0} \epsilon^{-1} (\theta(\vec{x}, \vec{w} + \epsilon(\delta_i - \vec{w})) - \theta(\vec{x}, \vec{w})) \tag{2}$$

where δ_i is the vector with 1 in position i and 0 elsewhere. These values are the empirical influence function or infinitesimal jackknife (Efron (1982)). Approximations for these values, or similar values that may be used in their place, are discussed in (Hesterberg (1995a)).

The x_i and w_i are sorted in order of increasing L_i. The integer and fractional parts M_i^I and M_i^F are computed as in Method 3, and x_i is again deterministically included in each created sample M_i^I times, but the random parts are generated using a type of stratified sampling. Let $M^F = \sum M_i^F$. M^F random uniform deviates are generated, one on (0,1), one on (1,2), ..., and the last on $(M^F - 1, M^F)$. The ith observation is chosen for the sample when a uniform deviate falls between $\sum_{j=1}^{i-1} M_j^F$ and $\sum_{j=1}^{i} M_j^F$. This is essentially giving the ith observation a "target" region with length proportional to M_i^F and sorted such that the observations with low L_i values appear together (first), while the observations with high L_i values appear together at the end. The ith observation is randomly included in the sample between 0 and 2 times (its target region may straddle the regions for two of the random variates), and the expected number of inclusions is M_i^F. Furthermore, the *cumulative* random frequencies closely match their expected frequencies.

The random part of the procedure is repeated K times (independently) and the final estimate has the same form as (1).

This procedure could be modified to introduce dependence and negative correlation between the K samples, but we have not done so. The simulation results below for Methods 4 and 5 suggest that this would be promising. The current Method 5 uses a total of KM^F random deviates, K each from each of the intervals $(0,1),(1,2),\ldots,(M^F-1,M^F)$. Each such interval could be partitioned into K equal subintervals, and the corresponding K deviates generated one from each of the subintervals.

These five methods for generating the pseudo-samples are summarized below, for easy reference:

- Method 1: include the ith observation round$(w_i \cdot M)$ times (deterministic)
- Method 2: form sample of size M by repeated sampling with replacement; the ith observation is chosen with probability proportional to w_i (simplest random method)
- Method 3: split Mw_i into integer and fractional part; append deterministic integer part and form a random sample using the fractional part.
- Method 4: similar to Method 3, except that all random parts are generated simultaneously, using a second decomposition into integer and fractional parts.
- Method 5: same as Method 3, except the random sampling is stratified after sorting by empirical influence function values (L_i).

3 Simulation Results

We tested these methods in a simulation study using two statistics: the univariate mean and the correlation coefficient. These statistics were chosen because there are weighted versions available for comparing to the estimates, and because the mean is linear while the correlation coefficient is quite nonlinear.

For each statistic, and for sample sizes $n = 10, 20, 40$ and 80, we generated 400 random data sets. For each data set we calculated the empirical influence function values L and a meaningful vector of weights (the weights used in one-sided bootstrap tilting confidence intervals with $\alpha = 0.025$), and calculated the five approximations, using $M = n, 3n, 10n$, and $30n$ and $K = 1$ and 20. We summarize the means and standard deviations of the errors $\hat{\theta} - \theta_w$, where θ_w is the desired value $\theta(\vec{x}, \vec{w})$ and $\hat{\theta}$ is the approximation. Note that the standard deviations include two components of variance — the variability given a set of random data and weights, and the variability between such samples. We also report the associated t-statistics to judge the bias of estimates.

We exclude results for Method 4 with $K = 1$, as this method is equivalent to Method 3 with $K = 1$. There is no K for Method 1, as the method is deterministic (conditional on the data and weights).

For the sample mean case, the data were generated from the univariate standard normal distribution. Results are listed in Tables 1-5. Method 1 (Table 1) is clearly biased when $M = n$; for larger M it is still biased but not enough to be apparent in the t-statistics, except with $n = 80$ and $M = 240$. The random methods (Methods 2-5, Tables 2-5) had low t-statistics, indicating no evidence of bias (in fact the methods are unbiased for the sample mean).

With $K = 1$, Method 5 performs extremely well, typically outperforming Method 3 by a factor of 10 in terms of error standard deviation (or 100 in terms of error variance). With $K = 20$, Method 4 is best for $n = 10$, but Method 5 is better for larger n.

For the correlation case, the data were bivariate normal with correlation $\sqrt{.5}$. Results are listed in Tables 6-10. Method 1 is again severely biased when $M = n$. The random methods (Tables 7-10) exhibit a downward bias (almost all the t-statistics are negative); this is not too surprising, as the correlation is a nonlinear statistic. The bias is substantial with $M = n$ but decreases rapidly as M increases.

With $K = 1$ Method 5 performs very well, in terms of both mean and standard deviation. With $K = 20$, Method 4 sometimes beats Method 5 (for small n and large M), but Method 5 is always competitive.

These conclusions are confirmed by examination of root mean square errors. (See Tables 11-14.)

4 Conclusion

Based on these results, we recommend Method 5, if the values L_j are available. It is always competitive with other methods, and sometimes substantially better. Otherwise Method 4 (which is equivalent to Method 3 if $K = 1$) is recommended.

We obtain more accurate results with larger values of M and smaller K. For example, in Table 10 (correlation, Method 5) with $n = 20$, the estimated standard deviation of the errors is .013 with $M = 20$ and $K = 20$ (400 total observations) and is only 0.0018 with $M = 600$ and $K = 1$ (600 total observations). Hence, if the statistic is relatively fast to compute for large sample sizes, it is best to let $K = 1$ and choose the sample size M as large as possible. However, if a computational time is proportional to M^2, M^3, or even exponential in M, then a smaller M and larger K would be appropriate.

These methods should be tested with other statistics.

References

Efron, B. (1981). Nonparametric Standard Errors and Confidence Intervals. *Canadian Journal of Statistics*, 9:139 – 172.

Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans.* National Science Foundation – Conference Board of the Mathematical Sciences Monograph 38. Society for Industrial and Applied Mathematics, Philadelphia.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap.* Chapman and Hall.

Hesterberg, T. C. (1995a). Tail-Specific Linear Approximations for Efficient Bootstrap Simulations. *Journal of Computational and Graphical Statistics*, 4(2):113–133.

Hesterberg, T. C. (1995b). Weighted Average Importance Sampling and Defensive Mixture Distributions. *Technometrics*, 37(2):185–194.

Table 1: Sample Mean Results, using approximation Method 1

n	θ_w mean	θ_w $\hat{\sigma}$	M	$\hat{\theta} - \theta_w$ mean	$\hat{\theta} - \theta_w$ $\hat{\sigma}$	$\hat{\theta} - \theta_w$ t
10	−0.58	0.34	10	−0.059	0.11	−10.6
			30	−0.001	0.035	−0.7
			100	−0.0003	0.0099	−0.6
			300	−0.0002	0.0035	−1.2
20	−0.44	0.23	20	0.057	0.092	12.3
			60	0.0004	0.022	0.4
			200	0.00005	0.0070	0.1
			600	−0.0002	0.0023	−1.5
40	−0.32	0.16	40	0.130	0.048	53.9
			120	0.0007	0.016	0.9
			400	−0.0002	0.0050	−0.7
			1200	−0.00002	0.0015	−0.3
80	−0.21	0.11	80	0.156	0.029	106.5
			240	−0.0022	0.010	−4.5
			800	0.0001	0.0031	0.9
			2400	−0.00003	0.0010	−0.7

The number of digits for each entry reflects the standard error for that entry—standard errors are the same order of magnitude as the last digit shown. Exception: t-statistics are all rounded to one decimal place.

Table 2: Sample Mean Results, using approximation Method 2

	θ_w			$\hat{\theta} - \theta_w$			
n	mean	$\hat{\sigma}$	M	mean	$\hat{\sigma}_{K=1}$	$\hat{\sigma}_{K=20}$	t
10	−0.58	0.34	10	−0.0005	0.28	0.064	−0.1
			30	−0.004	0.16	0.032	−2.3
			100	0.002	0.088	0.020	1.6
			300	0.0010	0.050	0.012	1.8
20	−0.44	0.23	20	0.002	0.21	0.047	0.7
			60	0.002	0.12	0.027	1.1
			200	0.0006	0.068	0.015	0.7
			600	−0.0002	0.039	0.0091	−0.4
40	−0.32	0.16	40	0.003	0.16	0.038	1.5
			120	0.000005	0.091	0.020	0.0
			400	0.0003	0.049	0.011	0.6
			1200	0.0004	0.029	0.0064	1.1
80	−0.21	0.11	80	−0.002	0.11	0.025	−1.4
			240	0.00006	0.064	0.015	0.1
			800	0.0003	0.035	0.0082	0.7
			2400	0.00005	0.020	0.0047	0.2

The t-statistics are calculated using $\hat{\sigma}_{K=1}$ with a sample size of 8000 $(400 \cdot 20)$. See Table 1 for information on rounding.

Table 3: Sample Mean Results, using approximation Method 3

n	θ_w		M	$\hat{\theta} - \theta_w$			
	mean	$\hat{\sigma}$		mean	$\hat{\sigma}_{K=1}$	$\hat{\sigma}_{K=20}$	t
10	-0.58	0.34	10	-0.002	0.19	0.042	-1.2
			30	-0.0007	0.070	0.016	-0.9
			100	0.0004	0.021	0.0048	1.8
			300	-0.00010	0.0071	0.0015	-1.4
20	-0.44	0.23	20	0.0007	0.15	0.031	0.4
			60	-0.0004	0.051	0.011	-0.9
			200	-0.00005	0.015	0.0035	-0.3
			600	-0.00006	0.0051	0.0011	-1.0
40	-0.32	0.16	40	0.001	0.11	0.025	0.9
			120	-0.0001	0.037	0.0087	-0.3
			400	-0.00008	0.011	0.0026	-0.7
			1200	0.00004	0.0037	0.00080	0.8
80	-0.21	0.11	80	0.0007	0.078	0.017	0.8
			240	0.00009	0.026	0.0060	0.3
			800	-0.00010	0.0078	0.0017	-1.1
			2400	0.000003	0.0026	0.00060	0.1

The t-statistics are calculated using $\hat{\sigma}_{K=1}$ with a sample size of 8000 $(400 \cdot 20)$. See Table 1 for information on rounding.

Table 4: Sample Mean Results, using approximation Method 4

	θ_w			$\hat{\theta} - \theta_w$		
n	mean	$\hat{\sigma}$	M	mean	$\hat{\sigma}_{K=20}$	t
10	-0.58	0.34	10	0.0004	0.010	0.8
			30	-0.00001	0.0034	-0.1
			100	-0.00005	0.0010	-1.0
			300	0.00002	0.00034	1.0
20	-0.44	0.23	20	-0.0001	0.0076	-0.3
			60	0.0002	0.0026	1.8
			200	0.00002	0.00076	0.5
			600	0.00002	0.00026	1.6
40	-0.32	0.16	40	-0.0003	0.0052	-1.3
			120	-0.00017	0.0019	-1.8
			400	0.00003	0.00055	1.0
			1200	-0.000001	0.00018	-0.1
80	-0.21	0.11	80	0.00002	0.0039	0.1
			240	0.00005	0.0013	0.7
			800	0.0000002	0.00039	0.0
			2400	0.000004	0.00014	0.6

The t-statistics are calculated using $\hat{\sigma}_{K=20}$ with a sample size of 400 (these are each the average of $K = 20$ dependent values).

Table 5: Sample Mean Results, using approximation Method 5

	θ_w			$\hat{\theta} - \theta_w$			
n	mean	$\hat{\sigma}$	M	mean	$\hat{\sigma}_{K=1}$	$\hat{\sigma}_{K=20}$	t
10	−0.58	0.34	10	−0.0013	0.067	0.015	−1.8
			30	−0.00004	0.022	0.0049	−0.2
			100	−0.00007	0.0066	0.0014	−0.9
			300	0.00002	0.0022	0.00049	0.7
20	−0.44	0.23	20	0.0003	0.032	0.0074	0.8
			60	−0.00004	0.0097	0.0023	−0.4
			200	0.00003	0.0030	0.00069	0.9
			600	0.00002	0.0010	0.00022	1.8
40	−0.32	0.16	40	0.00002	0.013	0.0030	0.2
			120	0.00004	0.0045	0.0010	0.7
			400	−0.00002	0.0013	0.00029	−1.4
			1200	0.000002	0.00044	0.00010	0.5
80	−0.21	0.11	80	−0.00002	0.0061	0.0014	−0.3
			240	−0.000005	0.0021	0.00044	−0.2
			800	0.000002	0.00060	0.00013	0.3
			2400	−0.000005	0.00020	0.000043	−2.3

The t-statistics are calculated using $\hat{\sigma}_{K=1}$ with a sample size of 8000 $(400 \cdot 20)$. See Table 1 for information on rounding.

Table 6: Sample Correlation Results, using approximation Method 1

	θ_w			$\hat{\theta} - \theta_w$		
n	mean	$\hat{\sigma}$	M	mean	$\hat{\sigma}$	t
10	0.37	0.31	10	−0.080	0.13	−12.2
			30	−0.006	0.043	−3.0
			100	−0.0001	0.015	−0.2
			300	0.0003	0.0041	1.3
20	0.44	0.20	20	0.004	0.064	1.3
			60	−0.002	0.023	−1.6
			200	0.0003	0.0064	0.8
			600	0.0002	0.0021	2.2
40	0.52	0.13	40	0.039	0.039	20.3
			120	0.0018	0.012	3.0
			400	−0.00007	0.0037	−0.4
			1200	0.00008	0.0012	1.4
80	0.58	0.075	80	0.055	0.021	52.2
			240	0.0023	0.0077	6.0
			800	−0.000003	0.0023	0.0
			2400	0.00003	0.00071	0.7

See Table 1 for information on rounding.

Table 7: Sample Correlation Results, using approximation Method 2

	θ_w			$\hat{\theta} - \theta_w$			
n	mean	$\hat{\sigma}$	M	mean	$\hat{\sigma}_{K=1}$	$\hat{\sigma}_{K=20}$	t
10	0.37	0.31	10	−0.047	0.30	0.078	−14.2
			30	−0.015	0.14	0.034	−9.5
			100	−0.0036	0.074	0.016	−4.3
			300	−0.0022	0.043	0.010	−4.6
20	0.44	0.20	20	−0.019	0.18	0.045	−9.2
			60	−0.004	0.096	0.022	−3.8
			200	−0.0023	0.050	0.011	−4.0
			600	−0.0009	0.030	0.0063	−2.9
40	0.52	0.13	40	−0.005	0.11	0.023	−4.2
			120	−0.0013	0.063	0.014	−1.8
			400	0.0005	0.034	0.0080	1.3
			1200	−0.0002	0.020	0.0044	−1.1
80	0.58	0.075	80	−0.0031	0.073	0.017	−3.8
			240	−0.0012	0.042	0.0099	−2.6
			800	−0.0005	0.023	0.0052	−2.0
			2400	−0.0003	0.013	0.0029	−2.0

The t-statistics are calculated using $\hat{\sigma}_{K=1}$ with a sample size of 8000 ($400 \cdot 20$). See Table 1 for information on rounding.

Table 8: Sample Correlation Results, using approximation Method 3

n	θ_w mean	θ_w $\hat{\sigma}$	M	$\hat{\theta} - \theta_w$ mean	$\hat{\sigma}_{K=1}$	$\hat{\sigma}_{K=20}$	t
10	0.37	0.31	10	−0.064	0.23	0.068	−24.6
			30	−0.0101	0.084	0.023	−10.7
			100	−0.0015	0.027	0.0064	−5.0
			300	−0.0001	0.0095	0.0021	−0.9
20	0.44	0.20	20	−0.026	0.13	0.035	−17.5
			60	−0.0032	0.048	0.011	−6.1
			200	−0.0003	0.015	0.0034	−1.9
			600	−0.00004	0.0050	0.0010	−0.6
40	0.52	0.13	40	−0.0114	0.083	0.018	−12.3
			120	−0.0009	0.028	0.0066	−2.7
			400	−0.00023	0.0086	0.0020	−2.4
			1200	−0.000005	0.0029	0.00062	−0.2
80	0.58	0.075	80	−0.0055	0.054	0.012	−9.2
			240	−0.0006	0.017	0.0037	−2.9
			800	−0.00006	0.0053	0.0012	−1.0
			2400	0.00001	0.0018	0.00043	0.6

The t-statistics are calculated using $\hat{\sigma}_{K=1}$ with a sample size of 8000 $(400 \cdot 20)$. See Table 1 for information on rounding.

Table 9: Sample Correlation Results, using approximation Method 4

	θ_w			$\hat{\theta} - \theta_w$		
n	mean	$\hat{\sigma}$	M	mean	$\hat{\sigma}_{K=20}$	t
10	0.37	0.31	10	-0.058	0.053	-22.0
			30	-0.0100	0.012	-16.5
			100	-0.0011	0.0027	-8.3
			300	-0.00014	0.00051	-5.4
20	0.44	0.20	20	-0.0246	0.018	-27.5
			60	-0.0033	0.0041	-15.8
			200	-0.00028	0.00087	-6.6
			600	-0.00004	0.00026	-3.8
40	0.52	0.13	40	-0.0115	0.0083	-27.6
			120	-0.00115	0.0016	-14.0
			400	-0.00009	0.00041	-4.2
			1200	-0.000009	0.00015	-1.1
80	0.58	0.075	80	-0.0047	0.0037	-25.5
			240	-0.00046	0.00096	-9.5
			800	-0.00002	0.00025	-1.5
			2400	-0.0000008	0.000090	-0.2

The t-statistics are calculated using $\hat{\sigma}_{K=20}$ with a sample size of 400 (these are each the average of $K = 20$ dependent values).

Table 10: Sample Correlation Results, using approximation Method 5

n	θ_w mean	θ_w $\hat{\sigma}$	M	$\hat{\theta} - \theta_w$ mean	$\hat{\sigma}_{K=1}$	$\hat{\sigma}_{K=20}$	t
10	0.37	0.31	10	−0.026	0.13	0.041	−18.5
			30	−0.0028	0.045	0.011	−5.4
			100	−0.0004	0.015	0.0035	−2.1
			300	0.00003	0.0047	0.0011	0.6
20	0.44	0.20	20	−0.0073	0.050	0.013	−12.9
			60	−0.0005	0.019	0.0039	−2.5
			200	−0.00011	0.0054	0.0013	−1.9
			600	−0.000003	0.0018	0.00034	−0.1
40	0.52	0.13	40	−0.0019	0.022	0.0050	−7.9
			120	−0.00020	0.0074	0.0017	−2.3
			400	−0.00002	0.0021	0.00045	−1.1
			1200	0.000003	0.00073	0.00018	0.3
80	0.58	0.075	80	−0.0005	0.0097	0.0023	−4.5
			240	0.00001	0.0030	0.00069	0.4
			800	0.000002	0.00093	0.00021	0.2
			2400	0.000002	0.00031	0.000074	0.5

The t-statistics are calculated using $\hat{\sigma}_{K=1}$ with a sample size of 8000 $(400 \cdot 20)$. See Table 1 for information on rounding.

Table 11: Summary of Sample Mean Results ($K = 1$)

		r.m.s. error of $(\hat{\theta} - \theta_w)$			
n	M	Method 1	Method 2	Method 3	Method 5
10	10	0.13	0.28	0.19	0.067
	30	0.035	0.16	0.070	0.022
	100	0.0099	0.088	0.021	0.0066
	300	0.0035	0.050	0.0071	0.0022
20	20	0.11	0.21	0.15	0.032
	60	0.022	0.12	0.051	0.0097
	200	0.0070	0.068	0.015	0.0030
	600	0.0023	0.039	0.0051	0.0010
40	40	0.14	0.16	0.11	0.013
	120	0.016	0.091	0.037	0.0045
	400	0.0050	0.050	0.011	0.0013
	1200	0.0015	0.029	0.0037	0.00044
80	80	0.16	0.11	0.078	0.0061
	240	0.010	0.064	0.026	0.0021
	800	0.0031	0.035	0.0078	0.00060
	2400	0.0010	0.020	0.0026	0.00020

Table 12: Summary of Sample Mean Results ($K = 20$)

n	M	r.m.s. error of $(\hat{\theta} - \theta_w)$			
		Method 2	Method 3	Method 4	Method 5
10	10	0.064	0.042	0.010	0.015
	30	0.033	0.016	0.0034	0.0049
	100	0.020	0.0048	0.0011	0.0014
	300	0.012	0.0015	0.00034	0.00049
20	20	0.047	0.031	0.0076	0.0074
	60	0.027	0.011	0.0026	0.0023
	200	0.015	0.0035	0.00076	0.00069
	600	0.0091	0.0011	0.00026	0.00022
40	40	0.038	0.025	0.0052	0.0030
	120	0.020	0.0087	0.0019	0.0010
	400	0.011	0.0026	0.00055	0.00029
	1200	0.0064	0.00080	0.00018	0.00010
80	80	0.025	0.017	0.0039	0.0014
	240	0.015	0.0060	0.0013	0.00044
	800	0.0082	0.0017	0.00039	0.00013
	2400	0.0047	0.00060	0.00014	0.00043

Table 13: Summary of Sample Correlation Results ($K = 1$)

n	M	r.m.s. error of $(\hat{\theta} - \theta_w)$			
		Method 1	Method 2	Method 3	Method 5
10	10	0.15	0.30	0.24	0.13
	30	0.044	0.15	0.085	0.045
	100	0.015	0.074	0.027	0.015
	300	0.0041	0.043	0.0095	0.0047
20	20	0.064	0.18	0.14	0.051
	60	0.023	0.096	0.048	0.019
	200	0.0064	0.050	0.015	0.0054
	600	0.0021	0.030	0.0050	0.0018
40	40	0.055	0.11	0.084	0.022
	120	0.012	0.063	0.028	0.0074
	400	0.0037	0.034	0.0086	0.0021
	1200	0.0012	0.020	0.0029	0.00073
80	80	0.059	0.073	0.054	0.0097
	240	0.0080	0.042	0.017	0.0030
	800	0.0023	0.023	0.0053	0.00093
	2400	0.00071	0.013	0.0018	0.00031

Table 14: Summary of Sample Correlation Results ($K = 20$)

n	M	r.m.s. error of $(\hat{\theta} - \theta_w)$			
		Method 2	Method 3	Method 4	Method 5
10	10	0.091	0.093	0.078	0.049
	30	0.038	0.025	0.016	0.011
	100	0.017	0.0065	0.0029	0.0035
	300	0.010	0.0021	0.00052	0.0012
20	20	0.049	0.044	0.030	0.015
	60	0.022	0.012	0.0052	0.0039
	200	0.012	0.0034	0.00091	0.0013
	600	0.0063	0.0010	0.00026	0.00034
40	40	0.024	0.022	0.014	0.0054
	120	0.014	0.0066	0.0020	0.0017
	400	0.0080	0.0020	0.00042	0.00045
	1200	0.0044	0.00062	0.00015	0.00018
80	80	0.017	0.013	0.0060	0.0024
	240	0.010	0.0037	0.0011	0.00069
	800	0.0052	0.0012	0.00025	0.00021
	2400	0.0029	0.00043	0.000090	0.00075