Jl. of Computers in Mathematics and Science Teaching (1993) 12(1), 83–92

The Introduction of Bootstrapping Techniques for Finding Confidence Intervals Using Simulation

KRISTIN KENNEDY AND PHYLLIS SCHUMACHER
Department of Mathematics
Bryant College, Smithfield, RI 02917, USA

The following paper details how the use of simulation can help to introduce computer intensive applications, namely bootstrapping, to a first year statistics course. The bootstrapping technique is used in creating confidence intervals for the mean, median, and variance. These bootstrap confidence intervals are compared to the traditional confidence intervals for the purpose of analyzing the accuracy of the bootstrapping technique.

INTRODUCTION

The use of simulation as a teaching technique in a statistics course has been discussed in previous papers, Olinsky and Schumacher (1990), and Kennedy, Olinsky, and Schumacher (1990). As observed in these papers, the hands-on use of a computer and the observation of sampling results through simulation provide the students with an intuitive understanding of complicated theoretical results. Simulation may also be used to demonstrate to students the effectiveness of nontraditional statistical procedures such as the computer intensive resampling procedures of which the best known is bootstrapping.

Bootstrapping is a statistical algorithmic procedure that may be used to calculate confidence interval estimates for unknown population parameters. The theoretical advantage of the bootstrapping method of calculating confidence intervals is that the population need not be normal.

As Stephen Turner observed in his paper "Elementary Bootstrapping with *Minitab*" (1988), the content of elementary statistics courses, particularly in business schools, is shifting towards computer-assisted data analysis. This shift away from the traditional hypothesis testing procedures allows for the examination of real data sets. The bootstrapping technique, because it has no model assumptions and is an algorithmic technique, is a likely addition to any statistics course, especially a data-analysis type course. The use of the computer and an easy to use statistical package such as *Minitab* are essential to the introduction of the bootstrapping algorithm in an elementary course.

Turner presents the bootstrapping procedure and provides a simple *Minitab* program to calculate a bootstrap confidence interval for the population mean. The following paper provides an illustration of the use of simulation to introduce bootstrapping in a basic statistics course. The simulation also emphasizes the meaning of traditional confidence intervals and compares the accuracy of these two techniques for the mean, median, and variance of a normal distribution. The programs were written in *Minitab* in order that they could be carried out with hands-on input from the students because *Minitab* is a popular statistics package.

To obtain a bootstrap interval for a population parameter, one starts with a random sample from the population for which no assumptions need be made. One then samples with replacement from this sample. Turner reports that 100 samples is considered adequate for the procedure. The statistic that is used to estimate the unknown parameter is then calculated for each of these 100 samples. The 100 statistics are then sorted and the appropriate percentiles chosen to obtain the desired interval.

We introduced simulation and bootstrapping in an elementary business statistics course and found that these concepts can be easily understood by elementary statistics students. The addition of this topic therefore enhanced the course by providing a method of calculating confidence intervals that is the same for all parameters, while giving the students a deeper understanding of the traditional methods.

DESIGN OF THE SIMULATION

The simulation was designed to illustrate the bootstrapping technique itself and to test and compare the accuracy of a 95% confidence interval for descriptive statistics, using both bootstrapping techniques and traditional methods. The entire *Minitab* program is found in the appendix. To create the initial random sample, we had to assume a certain distribution. Because most first-semester statistics courses rely heavily on the normal distribution, we calculated the 95% confidence intervals for a normal distribution with a mean = 0 and a variance = 1, and the program was run for a small sample size of n = 10 and a sample size of n = 25. The lack of normality is not a problem for large sample sizes; therefore, we used samples with a small size.

The program first calculates the traditional confidence intervals, which in itself was interesting to the students, because they could see how the different confidence intervals for each statistic were found using the *Minitab* program. The Minitab command, which can be found in line 9 of the program, for finding the t-interval of the sample in column C1 is MTB > TINTERVAL C1. Thus, the confidence interval for the mean, which assumes normality when $n < 30$, is obtained by a one-line call for a t-interval in the *Minitab* program. The confidence interval for the variance also assumes normality, but the calculation was a bit more complicated.

The formula for the confidence interval of the variance, which is included in many elementary texts because it contains the standard normal scores, is the large sample formula and is appropriate for sample sizes greater than 30. Because we were interested in smaller samples, we used the formula that is appropriate for all samples of size n from a normal population and requires the calculation of values from a chi-square distribution (Freund & Walpole, 1987). This formula for a sample of size n with a sample variance s^2 is as follows:

$$\frac{(n-1)s^2}{\chi^2_{(n-1,\ \frac{\alpha}{2})}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{(n-1,\ \frac{\alpha}{2})}}$$

The construction of the 95% confidence interval of the population variance is outlined in the *Minitab* subroutine VARIANCE.MTB called on line 15. It involves, for example, the inverse cumulative density function of the chi-square distribution, $\chi^2_{(n-1,\ .025)}$, that for a sample of size 10 is stored in the constant k3 in *Minitab* by the commands (see lines 18 through 24):

```
MTB > Let K1 = .025
MTB > Let K2 = 9
MTB > INVCDF K1 K3;
SUBC > CHISQUARE K2.
```

Normality is not an assumption for the traditional confidence interval of the median. In establishing an interval for the median, a table lookup was necessary. We used the tables from Inman and Conover whose text, *Modern Business Statistics* (1988), does include an optional section on calculating exact, small, sample confidence intervals for a population median. As noted by these authors, the level of confidence is not exactly 95% because of the discrete nature of the tabulated values. In particular, for samples of size 10 the actual percentage should be 97.85%, and for n=25 the actual percentage should be 95.67%, which is why the simulation results for our traditional 95% confidence intervals for the median did so well. It becomes obvious to the students how popular confidence intervals for the mean are because the others are more difficult to obtain. The interval for the mean was quite simple to find by a one-line call, but the other intervals were less accessible in *Minitab*.

Of the three intervals created, the only one that did not assume normality was the interval for the median. Bootstrapping is a technique that can be used to create a confidence interval for any parameter with no assumptions. The procedure always remains the same as opposed to the traditional methods that can vary greatly as we have shown. Bootstrap samples were created by selecting numbers from an original normal distribution with replacement. One hundred bootstrap samples each of size n = 10 or n = 25, depending on the experiment, were created. The mean, median, and variance were calculated for each of the 100 bootstrap samples. Those 100 values for each of the 3 statistics were stored in 3 different columns, and the columns were sorted. A 95% confidence interval for each statistic was built by averaging the 2nd and 3rd pieces of data in the sorted column, which would represent the lower bound of the interval and is in fact the 2.5 percentile, and by averaging the 97th and 98th pieces of data to represent the upper bound, which is the 97.5 percentile.

IMPLEMENTATION AND RESULTS

Because the student's background in *Minitab* was limited, and we did not want to lose valuable classroom time, we created the *Minitab* program, discussed it in class, and gave the program to each student. They in turn ran the program and returned to class with the data that could then be compiled together. The student's job was to run his or her program once and obtain 6 confidence intervals. We put two classes together so that we would have 100 groups of data. Thus each student brought back to class 6 confidence intervals. The first 3 would be the confidence intervals of the mean, median, and variance found by traditional methods, and the other 3 were those found by using bootstrapping methods. The class counted how many times out of 100 the given mean, median, and variance fell in the intervals for both methods and then calculated the percentages. This Monte Carlo approach clearly illustrated to the class what a statistician means by establishing a confidence interval for any statistic.

The information was placed in a table so that the results could be discussed in class. Table 1 shows the percentage of times the calculated statistic of the population actually fell within the confidence interval for $n = 10$. The levels of confidence using the traditional method for each of the three statistics performed accurately; however, in the bootstrapping method, the levels of confidence were less accurate. In fact the level of confidence for the variance was only 77%.

When the sample size was increased to $n = 25$, again the levels of confidence using the traditional method were as accurate as expected and are shown in Table 2. The bootstrapping method did perform better with a larger sample, but the level of confidence for the variance was still only 92%.

It is interesting to note that a small amount of accuracy may be sacrificed with the bootstrapping technique for the mean, but the variance sacrificed too much, especially for very small samples. Because the mean and the median are considered more stable statistics than the variance, we expected a loss of accuracy, but we were all surprised to see that the highest accuracy we attained in the confidence interval was only 92%. The results were also plotted on histograms using *Harvard Graphics™* that nicely indicated that the bootstrapping intervals gave results close to the traditional methods except for that of the variance.

In fact, only for the small sample of $n = 10$ was the loss of accuracy substantial. As an overhead presentation in class, these charts were a good illustration of the compiled results of all the students.

CONCLUSION

The educational values of this exercise are many. First of all the class sees concretely what a statistician means by a 95% confidence interval. By actually counting how many times the mean, median, and variance do fall within an expected range, a 95% confidence interval means then that only 2 or 3 misses are expected each time the program is run 100 times. Usually the student is surprised to actually see this fact through simulation.

The concept of bootstrapping is not covered in most first year textbooks, and it is a sampling technique that is not above the educational expertise of first year students. This simple program teaches the class the idea of bootstrapping, while illustrating that bootstrapping is not quite as accurate as traditional methods when working with data that is normal.

However, it also shows that bootstrapping is not terribly inaccurate. We also found that by showing them how to construct a bootstrap 95% confidence interval, that is, by constructing appropriate percentiles of sorted data, the student had a clearer idea of confidence intervals in the sense that he had a clearer idea of what a statistician means by the spread of data.

Students also learned that obtaining a confidence interval for the median and the variance in the traditional method was not simple. A table lookup was necessary for the median, and the chi-square distribution was needed for the variance.

Table 1
95% Confidence Intervals Containing the Parameter
Sample Size $n = 10$

Traditional Interval		*Bootstrap Interval*	
Parameter	Proportion	Parameter	Proportion
Mean	98/100	Mean	96/100
Median	99/100	Median	94/100
Variance	97/100	Variance	77/100

Table 2
95% Confidence Intervals Containing the Parameter
Sample Size $n = 25$

Traditional Interval		*Bootstrap Interval*	
Parameter	Proportion	Parameter	Proportion
Mean	95/100	Mean	94/100
Median	95/100	Median	94/100
Variance	95/100	Variance	92/100

One of the strengths of bootstrapping is that the technique of obtaining the confidence interval remains the same for any parameter and in fact for any distribution. However, the student must have access to a computer and a statistical package. The program is computer intensive; we can only use the bootstrapping technique with the aid of a computer. This program is a nice application of the use of computers in a first year statistics course, and theoretically the students can comprehend the results. Time, of course, is always an issue in any class, and every professor feels constricted by time. By giving the student the program, taking a few moments to discuss the idea of bootstrapping, and having the student run the program and return with the results, the educational benefits from the class discussions were well worth the time.

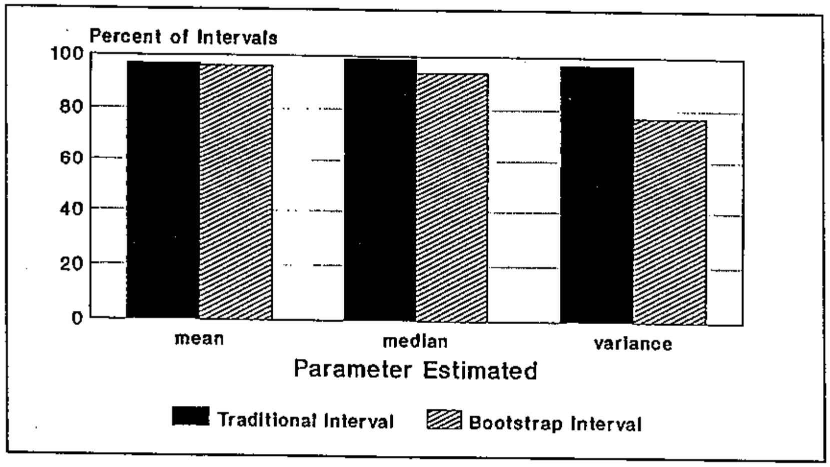

Figure 1. Sample intervals which contain parameter. Sample size n = 10.

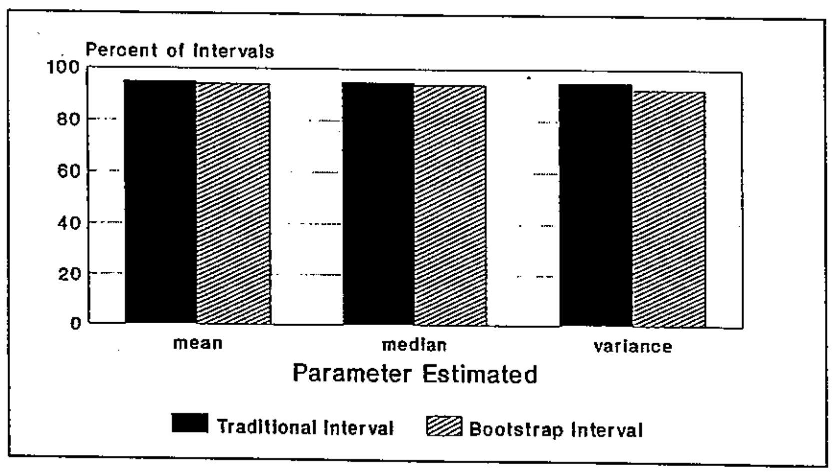

Figure 2. Sample intervals which contain parameter. Sample size $n = 25$.

References

Freund, J.E., & Walpole, R.E. (1987). *Mathematical statistics*. NJ: Prentice Hall.

Inman, R.L., & Conover, W.J. (1983). *Modern business statistics*. NY: John Wiley Sons.

Kennedy, K., Olinsky, A., & Schumacher, P. (1990). Using simulation as an integrated teaching tool in the mathematics classroom. *Education*, *3*(2).

Olinsky, A., & Schumacher, P. (1990). Illustration of regression results using simulation and graphical presentation. *Collegiate Microcomputer*, *8*(4).

Turner, S. (1988). Elementary bootstrapping with *Minitab*. *Collegiate Microcomputer*, *6*, 329–332.

APPENDIX

Minitab Program to Produce Confidence Intervals
Sample Size n = 10

	CODE	EXPLANATION
1.	OUTFILE 'Intervals'	Creates output file Intervals.Lis
2.	LET K50 = 10	Sets the sample size
3.	LET k60 = 2	
4.	LET k61 = 9	Percentiles used for Median Table
5.	EXECUTE 'CI'	Executes Main Subroutine CI.MTW
	CI.MTB	Command STORE 'CI' creates File
6.	NOECHO	Suppresses commands to screen
7.	OH=O	Suppresses page prompts
8.	RANDOM K50 C1	Places a random Normal (0, 1) sample of size n in Column 1
9.	TINTERVAL C1	Prints the traditional mean CI
10.	SORT C1 C600	Puts Sorted Data in Column 600
11.	LET C601 (1) =C600 (k60)	
12.	LET C601 (2) =C600 (k61)	Sets up median CI in Column 601
13.	NAME C601 'MED-CI'	
14.	PRINT C601	Prints the traditional median CI
15.	EXECUTE 'VARIANCE'	Calls subroutine for traditional Variance CI
16.	EXECUTE 'CILOOP'	Calls subroutine for Bootstrap CI's

```
17. END

    VARIANCE.MTB

18. LET K1=.025
19. LET K2=.975
20. LET K51=(K50-1)
21. INVCDF K1 K3;
22. CHISQUARE K51.                   18 thru 24 put 2,5th percentile of
                                     chi-square with n-1 degrees of
                                     freedom in k3 and 95th in k4
23. INVCDF K2 K4;
24. CHISQUARE 51.
25. LET K5 = VARIANCE (C1)
26. LET C500 (1)= ((K51*K5) /K4)
27. LET C500 (2)= ((K51*K5) /K3)
28. NAME C500 'VARCI'                Upper and Lower limits of the
                                     Traditional CI for Variance Found
29. PRINT C500
30. END

    CILOOP.MTB                       Subroutine Which creates
                                     Bootstrap sample and CI's

31. K1=2
32. EXECUTE 'BOOT' 100 Times         BOOT.MTB creates sample
33. LET K1=2
34. LET K2=1
35. EXECUTE 'MEAN' 100 Times         MEAN.MTB creates mean CI
36. SORT C200 C200
37. NAME C200 'means'
38. LET C201(1) = (C200(2) +
    C200(3))/2
39. LET C201(2) = (C200(97) +
    C200(98))/2
40. NAME C201 'CI-MEAN'
41. PRINT C201
42. LET K1=2
43. LET K2=1
44. EXECUTE 'VAR' 100 Times          VAR.MTB creates variance CI
45. SORT C300 C300
46. NAME C300 'vars'
47. LET C301(1) = (C300(2) +
    C300(3))/2
```

```
48. LET C301(2) = (C300(97) +
    C300(98))/2
49. NAME C301 'VAR-STD'
50. PRINT C301
51. LET K1=2
52. LET K2=1
53. EXECUTE 'MEDIAN' 100          MEDIAN.MTB creates median CI
    Times
54. SORT C400 C400
55. NAME C400 'medians'
56. LET C401(1) = (C400(2) +
    C400(3))/2
57. LET C401(2) = (C400(97) +
    C400(98))/2
58. NAME C401 'CI-MED'
59. PRINT C401
60. END

    BOOT.MTB
69. SAMPLE K50 C1 CK1;
70. REPLACE.
71. LET K1=K1+1
72. END

    MEAN.MTB
73. LET C200 (K2) =MEAN(CK1)
74. LET K1=K1+1
75. LET K2=K2+1
76. END

    VAR.MTB
76. LET C300(K2) = VARIANCE(CK1)
77. LET K1 = K1+1
78. LET K2 = K2+1
79. END

    MEDIAN.MTB
80. LET CK400(K2) = MEDIAN (Ck1)
81. LET K1 = K1+1
82. LET K2 = K2+1
83. END
```