

Scenario: When a shipment of coal is traded, the value of the shipment is determined by a number of its properties. One important property is the *gross calorific value*, which characterizes the heat content in megajoules per kilogram (MJ/kg). The International Organization of Standardization (ISO) issues a standard procedure (ISO 1928) for measuring the gross calorific value. When the procedure is carried out properly, resulting measurement errors are known to be approximately normal with a standard deviation of 0.1 MJ/kg. In the following table, 23 ISO 1928 measurements are given for a shipment of Osterfeld coal coded 262DE27:

23.640	23.678	23.690	23.700	23.712	23.727	23.730	23.740	23.760	23.771	23.778	23.780
23.796	23.800	23.830	23.840	23.850	23.860	23.860	23.870	23.877	23.890	23.940	

Source: A.M.H. van der Veen and A.J.M. Broos. Interlaboratory study programme "ILS coal characterization" - reported data. Technical report, NMI Van Swinden Laboratorium B.V., The Netherlands, 1996.

If you were to calculate the average of these 23 values, you would find the sample average to be 23.788.

1. Use parametric methods to construct and interpret a 95% confidence interval for μ .

95% CI: _____ Interpretation: _____

2. Use parametric methods to construct a 99% confidence interval for μ . Why is this 99% interval wider than the 95%?

99% CI: _____ Why it's wider: _____

3. Look again at the 23 values listed in the table (ordered from smallest to largest value). What proportion of these 23 values fall within your 95% confidence interval? Do we expect 95% of these values to fall within our 95% confidence interval? Explain.

Proportion of values that fall within 95% CI: _____

Explain why we do or do not expect 95% of observations to fall within a 95% confidence interval:

Scenario: Suppose we receive another shipment of 258GB41 coal from Daw Mill, where there were actually some doubts whether the stated accuracy of the ISO 1928 method was attained. We, therefore, doubt whether $\sigma = 0.1$ MJ/kg. We measure 22 ISO 1928 measurements for the shipment and find the values listed in the table to the right.

The sample average of these 22 values is 31.012 and the sample standard deviation is 0.1294.

ISO
30.800
30.810
30.830
30.880
30.920
30.921
30.929
30.990
30.990
30.991
31.020
31.024
31.026
31.030
31.050
31.060
31.060
31.091
31.125
31.170
31.208
31.330

4. Use parametric methods to construct a 95% confidence interval for μ . How does the width of this interval compare to the width of the interval you constructed in question #1? Explain why this interval is wider or more narrow than the interval you constructed in question #1.

95% CI: _____

Explain why this interval is wider or more narrow than the interval from question #1:

5. As we learned in class, these methods only apply if the normal distribution is a good model for our sample averages. For the confidence intervals you've constructed in this assignment, the sample sizes were $n=22$ and $n=23$. These sample sizes are not large enough for the Central Limit Theorem to kick-in. Why are we still confident that the normality assumption has been satisfied?

The normality assumption has been satisfied because: _____

6. Using the data on the top of this page, let's construct a 95% CI for μ using the *bootstrap method*. To do this, go to: http://lock5stat.com/statkey/bootstrap_1_quant/bootstrap_1_quant.html (or <http://goo.gl/nUX5eh>).

- Click the EDIT DATA button on the top
- Enter the data from the top of this page (paste it from the table at the top-right of this page)
- Generate at least 10,000 bootstrap samples by repeatedly clicking the GENERATE 1000 SAMPLES button
- Check the TWO-TAILED box and verify that 0.025 is being cut-off from each side of the distribution
- The confidence interval bounds (limits) are displayed in blue boxes below.
- Calculate the width and center of your bootstrap interval. Fill-in the following blanks to compare this bootstrap interval to the interval you constructed in question #4.

<u>Method</u>	<u>95% Confidence Interval</u>	<u>Width of interval</u>	<u>Center of interval</u>
Parametric	_____	_____	_____
Bootstrap	_____	_____	_____

Scenario: In January of 2002, the Euro was introduced. Soon after, stories started to circulate that some of the Euro coins would not be fair coins, because the "national side" of some coins would be too heavy or too light (see, for example, the New Scientist of January 4, 2002). A Belgian 1 Euro was tossed 250 times which resulted in 140 heads and 110 tails.

7. Construct two 95% confidence intervals for the proportion of heads this coin will yield. Use parametric methods to construct the first interval. Then, use bootstrap methods to construct the second 95% confidence interval. To do this, go to: http://lock5stat.com/statkey/bootstrap_1_cat/bootstrap_1_cat.html (or <http://goo.gl/0nafrl>)

- a) Click the EDIT DATA button on the top
- b) For this scenario, enter 140 for the count and 250 for the sample size
- c) Generate at least 10,000 bootstrap samples by repeatedly clicking the GENERATE 1000 SAMPLES button
- d) Check the TWO-TAILED box and verify that 0.025 is being cut-off from each side of the distribution
- e) The confidence interval bounds (limits) are displayed in blue boxes below.
- f) Calculate the width & center of your bootstrap interval. Fill-in the following blanks to compare your intervals.

<u>Method</u>	<u>95% Confidence Interval</u>	<u>Width of interval</u>	<u>Center of interval</u>
Parametric	_____	_____	_____
Bootstrap	_____	_____	_____

Scenario: In June of 2012, New York City Mayor Michael Bloomberg proposed banning the sale of soda or other sugary drinks in containers larger than 16 oz. A poll of 1093 randomly selected New York City voters conducted by Quinnipiac University taken shortly after the mayor's announcement found that 46.02% of them supported the ban.

8. Use parametric or bootstrap methods to construct a 95% confidence interval for the probability that a randomly selected New York City voter would support the ban. Briefly write out anything you can conclude from this interval.

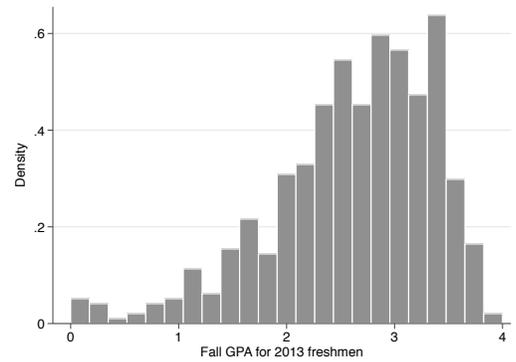
95% CI: _____ Conclusions: _____

9. Suppose ten times as many voters (10,930) were surveyed and again 46.02% were found to support the ban. Construct a 95% confidence interval and briefly explain why this interval is wider or more narrow than the interval you constructed in question #8.

95% CI: _____ Explanation: _____

Scenario: The average Fall semester GPA for $n=569$ freshmen at St. Ambrose in 2013 was 2.615 with a standard deviation of 0.7555. The distribution of these 569 GPAs, displayed to the right, shows a negative skew.

We're going to pretend this is our (unknown) population distribution.



10. Using a computer, I randomly selected 20 GPAs from this distribution. Those values are displayed in the table to the right.

The average of this sample of data is 2.472 with a standard deviation of 0.97356.

Construct 95% confidence intervals using parametric methods. To speed things up, you may want to use the confidence interval calculator at: <http://www.rossmanchance.com/applets/TBIA.html>. To do so, select ONE MEAN from the scenario menu. Then, you can either enter the summary information (n , mean, std. deviation) or paste the data. Finally, check the box at the right to construct a 95% interval.

95% CI: _____

GPA
1.00
1.18
1.22
1.42
1.57
1.58
1.73
1.83
2.09
2.15
2.80
3.08
3.29
3.33
3.37
3.41
3.47
3.56
3.61
3.75

11. Using the same data from that table to the right, construct a 95% confidence interval using bootstrap methods. To do this, once again go to:

http://lock5stat.com/statkey/bootstrap_1_quant/bootstrap_1_quant.html
(or <http://goo.gl/nUX5eh>) and:

- Click the EDIT DATA button on the top
- Enter the data (paste it from the table at on the right of this page)
- Generate at least 10,000 bootstrap samples by repeatedly clicking the GENERATE 1000 SAMPLES button
- Check the TWO-TAILED box and verify that 0.025 is being cut-off from each side of the distribution
- The confidence interval bounds (limits) are displayed in blue boxes below.
- Calculate the width and center of your bootstrap interval. Fill-in the following blanks to compare this bootstrap interval to the interval you constructed in question #10.

<u>Method</u>	<u>95% Confidence Interval</u>	<u>Width of interval</u>	<u>Center of interval</u>
Parametric	_____	_____	_____
Bootstrap	_____	_____	_____

12. Look at the intervals, widths, and centers you just wrote for the previous question. Now look at the original data from the scenario. We had a population mean of 2.615. Which confidence interval – the parametric or bootstrap – turned out to be the “better” interval. Explain your decision.

The better interval was the (circle one): PARAMETRIC BOOTSTRAP

Explanation: _____

13. Finally, enter the data once again into the applet you used to construct bootstrap confidence intervals. Construct a 95% confidence interval for the median Fall semester GPA.

95% CI for the population median: _____