Chapter 8
Inferences Based on a Single Sample: Tests of Hypothesis

8.1 Introduction

Chapter 8 introduces the reader to the topic of tests of hypothesis. Three parameters - the population mean, proportion, and variance - are studied in the chapter. **XLSTAT** provides techniques for conducting tests of hypothesis for the population mean and the population proportion, but does not offer a test for the population variance.

We find the tests of hypothesis in the **Parametric Tests** menu within the **XLSTAT** tab. **XLSTAT** offers both large and small tests for a population mean and a large sample test for a population proportion.

We note that the test of hypothesis material covered in *A First Course in Statistics* appears in Chapter 6 of that text.

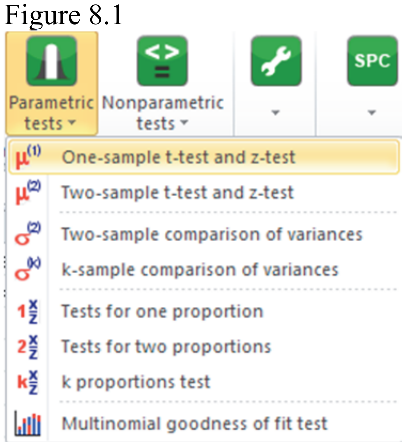
The following examples from *Statistics* illustrate the test of hypothesis calculations that can be found using XLSTAT in this chapter:

Excel Companion Exercise	Page	Statistics Example	Excel File Name
8.1	109	Example 8.4	DRUGRAT
8.2	112	Example 8.7	EMISSIONS
8.3	115	Example 8.9	

8.2 Test of Hypothesis for a Population Mean Using the Z-distribution

When conducting a test of hypothesis for a population mean, it is sometimes possible that the population standard deviation will be known. In such cases, it is possible to utilize the z-distribution in the calculation of the test statistic. To use this tool within **XLSTAT**, **open** a new workbook and place the cursor in the upper left cell of the worksheet. Click on the **XLSATAT Add-In** menu. Click on the **Parametric tests** option to access the **One-sample t-test and z-test** menu (see Figure 8.1).

Figure 8.1



We illustrate how to use this technique with the following exercise:

Exercise 8.1: We use Example 8.4 found in the *Statistics* text.

Problem The effect of drugs and alcohol on the nervous system has been the subject of considerable research. Suppose a research neurologist is testing the effect of a drug on response time by injecting 100 rats with a unit dose of the drug, subjecting each rat to a neurological stimulus, and recording its response time (data shown in Table 8.1 from the DRUGRAT data file). The neurologist knows that the mean response time for rats not injected with the drug (the "control" mean) is 1.2 seconds. Conduct a test of hypothesis to determine if the mean response time of the rats differs from 1.2 seconds. Use $\alpha = .01$.

Table 8.1

1.90	2.00	0.86	1.84	0.70	0.98	0.57	0.21	0.55	0.61
2.17	1.27	1.19	0.80	0.54	1.14	0.27	0.79	1.65	0.05
0.61	0.98	0.79	0.64	1.40	1.16	0.51	0.94	0.81	1.21
1.17	1.55	1.37	1.08	1.06	1.64	1.27	0.45	1.00	0.48
0.66	0.64	1.31	0.74	0.54	1.16	1.81	1.19	2.55	1.63
1.86	0.60	0.85	0.93	0.17	1.01	0.88	1.60	1.96	1.45
1.41	1.55	0.71	1.71	0.98	1.09	0.31	0.14	1.31	0.22
1.30	0.93	1.21	1.05	0.89	0.77	0.92	0.99	1.88	0.49
0.70	0.48	1.23	1.44	1.28	1.58	0.93	1.08	1.51	1.29
0.56	0.39	0.89	0.42	0.68	0.99	1.66	1.57	1.48	1.40

Solution:

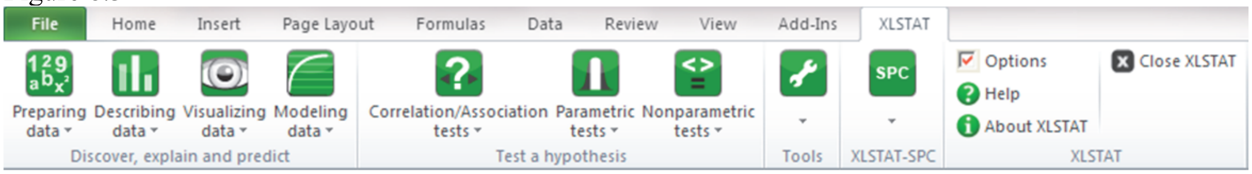
We solve Exercise 8.1 utilizing the **One-sample t-test and z-test** menu presented in **XLSTAT**. **Open** the data file **DRUGRAT** by following the directions found in the preface of this manual. If done correctly, the data should appear in a workbook similar to that shown in Figure 8.2.

Figure 8.2

A
TIME
1.9
2.17
0.61
1.17
0.66
1.86
1.41
1.3
0.7
0.56
2
1.27

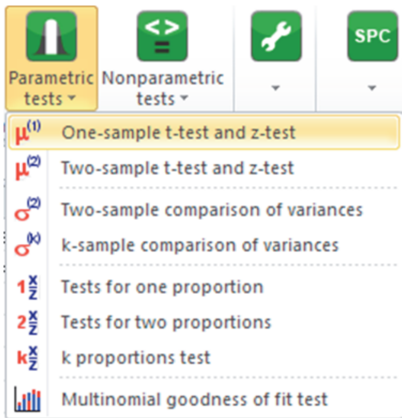
To conduct the desired test of hypothesis, we click on the **XLSTAT** tab at the top of the **Excel** workbook to access the **XLSTAT** menus shown in Figure 8.3.

Figure 8.3



To conduct the desired test of hypothesis, we click on the **Parametric tests** menu and select the **One-sample t-test and z-test** option shown in Figure 8.4.

Figure 8.4



This opens the **One-sample t-test and z-test** menu shown in Figures 8.5-8.6. We need to first specify the location of the data that is to be analyzed. In our data set, the data is located in Column A, rows 2 – 101, with row 1 being the variable label. We specify the location in the **Data** box and check the **Column labels** box to indicate the first row of data represents the variable name, **TIME**. Please note that you may choose to drag the mouse over the range of the data to be included instead of typing the location in the data box as previously described. In order to utilize the z-distribution in the calculation of the test statistic, we need to check the **z test** box in this menu.

Click on the **Options** tab (shown in Figure 8.6) to specify the **alternative hypothesis**, the **theoretical mean** (which is the hypothesized value of the mean we are testing in the null hypothesis), the **significance level**, and to also indicate how we want **XLSTAT** to work with the population variance required in the test statistic calculations. If this value is known, we check the **User defined** button and enter the value in the **Variance** box. If not, we check the **Estimated using samples** button. In our example, we have entered **Mean 1 ≠ Theoretical mean**, the value of **1.2** as the **Theoretical mean,** the value **1%** as the **Significance level,** and checked the **Estimated using samples** button. Click **OK** to conduct the test of hypothesis.

Figure 8.5

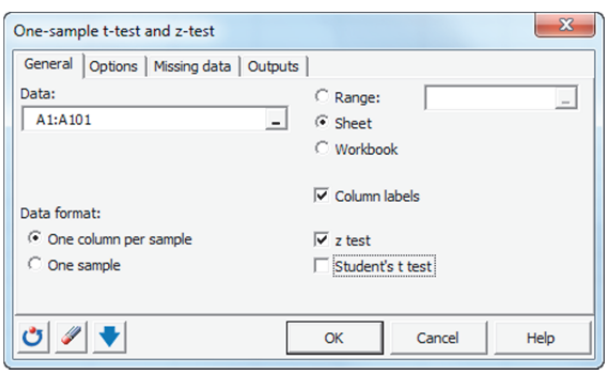
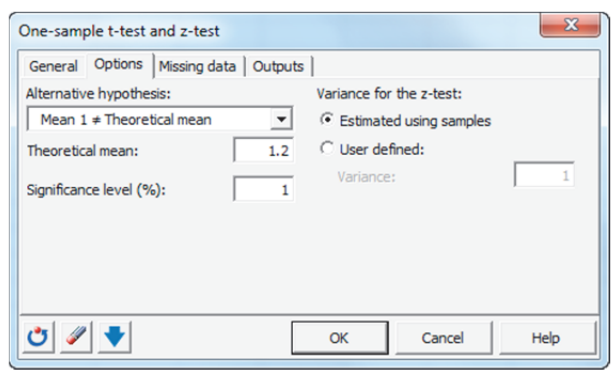


Figure 8.6



The XLSTAT output is shown in Figure 8.7.

Figure 8.7

Difference	-0.1483
z (Observed value)	-2.9766
|z| (Critical value)	2.5758
p-value (Two-tailed)	0.0029
alpha	0.01

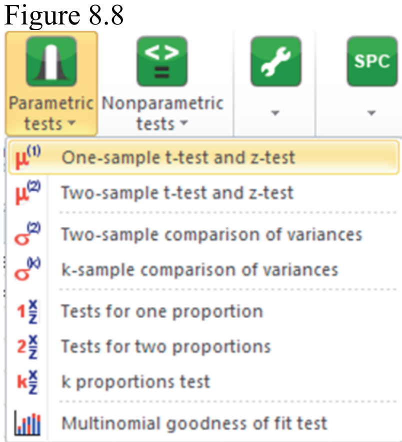
Test interpretation:
H0: The difference between the means is equal to 0.
Ha: The difference between the means is different from 0.
As the computed p-value is lower than the significance level alpha=0.01,
one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.
The risk to reject the null hypothesis H0 while it is true is lower than 0.29%.

We compare this printout to the information shown in the text. We note that the test statistic shown in the printout (z = -2.9766) has been rounded in the text (z = -3.0).

8.3 Test of Hypothesis for a Population Mean Using the t-distribution

When conducting a test of a population mean, it is likely that the value of the population standard deviation will be unknown. In such cases, the t-distribution is used in the calculation of the test statistic. To use the t-distribution test within **XLSTAT**, **open** a new workbook and place the cursor in the upper left cell of the worksheet. Click on the **XLSATAT Add-In** menu. Click on the **Parametric tests** option to access the **One-sample t-test and z-test** menu (see Figure 8.8).

Figure 8.8



We illustrate how to use this technique with the following exercise:

Exercise 8.2: We use Example 8.7 found in the *Statistics* text.

Problem A major car manufacturer wants to test a new engine to determine whether it meets new air-pollution standards. The mean emission μ for all engines of this type must be less than 20 parts per million of carbon. Ten engines are manufactured for testing purposes, and the emission level of each is determined. The data (in parts per million) are listed in Table 8.2 (from the EMISSIONS data file). Does the data supply sufficient evidence to allow the manufacturer to conclude that this type of engine meets the pollution standard? Assume that the manufacturer is willing to risk a Type I error with probability $\alpha = .01$.

Table 8.2

E-LEVEL									
15.6	16.2	22.5	20.5	16.4	19.4	19.6	17.9	12.7	14.9

Solution:

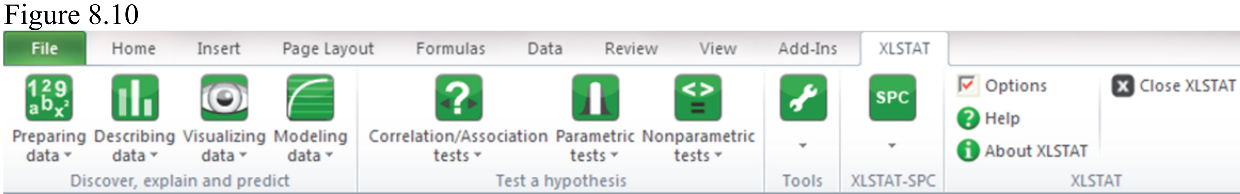
We solve Exercise 8.2 utilizing the **One-sample t-test and z-test** menu presented in **XLSTAT**. **Open** the data file **EMISSIONS** by following the directions found in the preface of this manual. If done correctly, the data should appear in a workbook similar to that shown in Figure 8.9.

Figure 8.9

	A
1	E-LEVEL
2	15.6
3	16.2
4	22.5
5	20.5
6	16.4
7	19.4
8	19.6
9	17.9
10	12.7
11	14.9

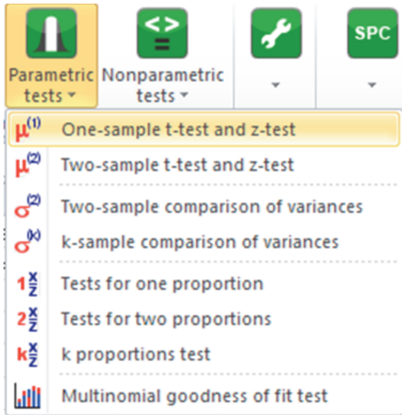
To conduct the desired test of hypothesis, we click on the **XLSTAT** tab at the top of the **Excel** workbook to access the **XLSTAT** menus shown in Figure 8.10.

Figure 8.10



To conduct the desired test of hypothesis, we click on the **Parametric tests** menu and select the **One-sample t-test and z-test** option shown in Figure 8.11.

Figure 8.11



This opens the **One-sample t-test and z-test** menu shown in Figures 8.12-8.13. We need to first specify the location of the data that is to be analyzed. In our data set, the data is located in Column A, rows 2 – 11, with row 1 being the variable label. We specify the location in the **Data** box and check the **Column labels** box to indicate the first row of data represents the variable name, **E-LEVEL**. Please note that you may choose to drag the mouse over the range of the data to be included instead of typing the location in the data box as previously described. In order to utilize the t-distribution in the calculation of the test statistic, we need to check the **Student's t test** box in this menu.

Click on the **Options** tab (shown in Figure 8.13) to specify the **alternative hypothesis**, the **theoretical mean** (which is the hypothesized value of the mean we are testing in the null hypothesis), and the **significance level**. In our example, we have entered **Mean 1 < Theoretical mean**, the value of **20** as the **Theoretical mean,** and the value **1%** as the **Significance level**. Click **OK** to conduct the test of hypothesis.

Figure 8.12

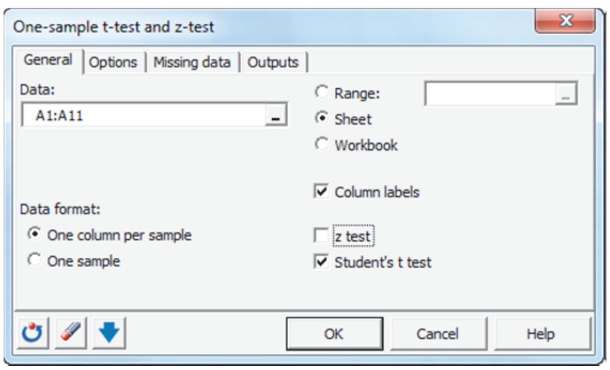
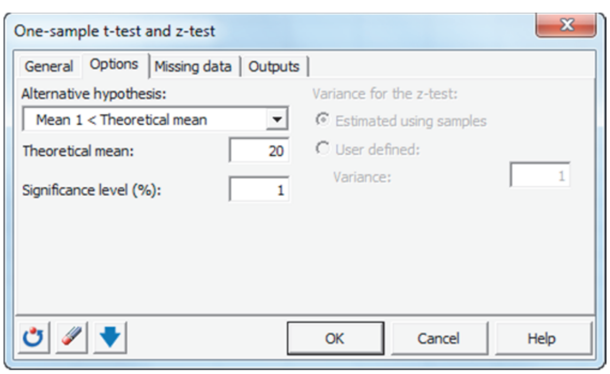


Figure 8.13



The XLSTAT output is shown in Figure 8.14.

Figure 8.14

Difference	-2.4300
t (Observed value)	-2.6029
t (Critical value)	-2.8234
DF	9
p-value (one-tailed)	0.0143
alpha	0.01

Test interpretation:

H0: The difference between the means is equal to 0.
Ha: The difference between the means is lower than 0.

As the computed p-value is greater than the significance level alpha=0.01, one cannot reject the null hypothesis H0.

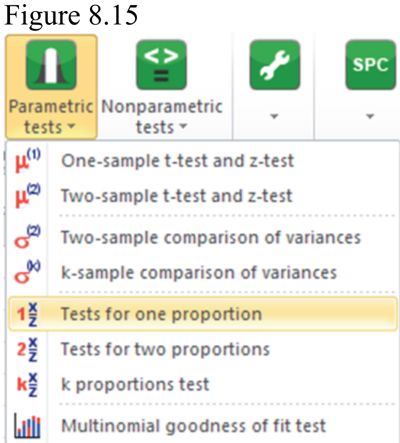
The risk to reject the null hypothesis H0 while it is true is 1.43%.

We compare this printout to the information shown in the text. We note that the test statistic and p-value shown in the printout are identical to the values shown in the text.

8.4 Tests of Hypothesis for a Population Proportion

When testing a population proportion, XLSTAT requires the user to specify the test they want to conduct. They must also enter the sample size, the number of successes, and the significance level of the test . To use the test of hypothesis tool within **XLSTAT**, **open** a new workbook and place the cursor in the upper left cell of the worksheet. Click on the **XLSATAT Add-In** menu. Click on the **Parametric tests** option to access the **Tests for one proportion** menu (see Figure 8.15).

Figure 8.15



We illustrate using the next exercise.

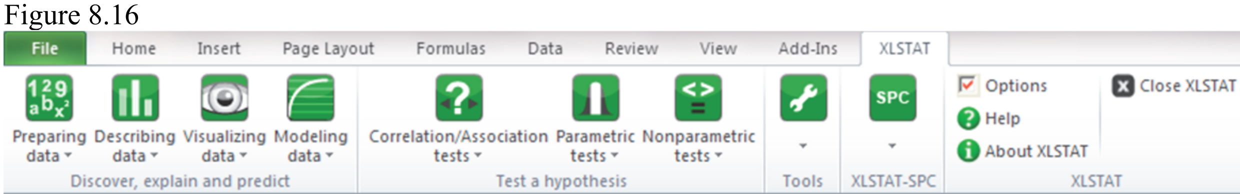
Exercise 8.3: We use Example 8.9 from the *Statistics* text.

Problem The reputations (and hence sales) of many businesses can be severely damaged by shipments of manufactured items that contain a large percentage of defectives. For example, a manufacturer of alkaline batteries may want to reasonably certain that less than 5% of its batteries are defective. Suppose 300 batteries are randomly selected from a very large shipment, each is tested, and 10 defective batteries are found. Does this outcome provide sufficient evidence for the manufacturer to conclude that the fraction defective in the entire shipment is less than .05? Use $\alpha = .01$.

Solution:

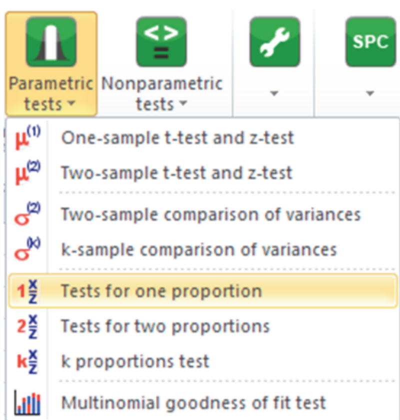
We solve Exercise 8.3 utilizing the **Tests for one proportion** menu presented in **XLSTAT**. To create the desired test, we click on the **XLSTAT** tab at the top of the **Excel** workbook to access the **XLSTAT** menus shown in Figure 8.16.

Figure 8.16



To generate the desired test, we click on the **Parametric tests** menu and select the **Tests for one proportion** option shown in Figure 8.17.

Figure 8.17



This opens the **Tests for one proportion** menu shown in Figures 8.18-8.19. We need to first select how the data is presented in the problem. If the data is given in terms of the number of successes, we click on the **Frequency** button. If the data is given in terms of the proportion of successes, we would click on the **Proportion** button. In this problem, we are told that 10 of the 300 batteries are defective. We click on the **Frequency** button and enter a **Frequency** of **10** and a **Sample size** of **300**. We also specify a **Test proportion:** of **.05**.

Click on the **Options** tab (shown in Figure 8.19) to specify the **Alternative hypothesis**, the **Hypothesized difference (D)**, and the **significance level** (which is the opposite of the confidence level). To test if the proportion of all batteries is less than .05, we test **Proportion – Test proportion < D** and type the value of the **Hypothesized difference (D)** of **0**. In our example, we have entered **1%** as the **Significance level (%)**. We also make sure to click on the **Wald** button to generate the test of hypothesis we desire. Click **OK** to conduct the test.

Figure 8.18

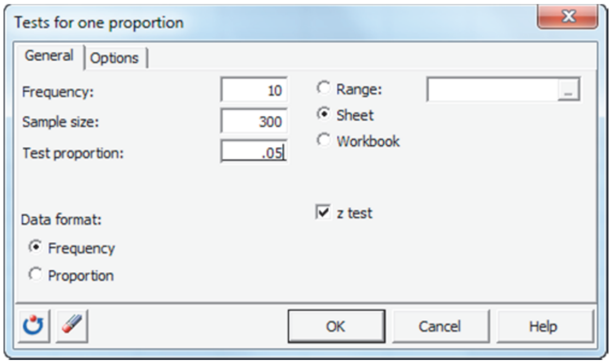
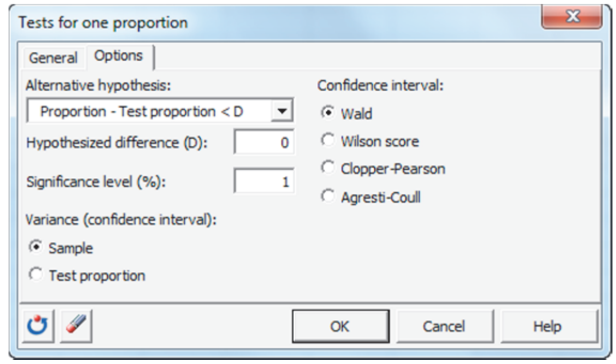


Figure 8.19



The **XLSTAT** output is shown in Figure 8.20.

Figure 8.20

z-test for one proportion / Lower-tailed test:

Difference	-0.0167
z (Observed value)	-1.3245
z (Critical value)	-2.3263
p-value (one-tailed)	0.0927
alpha	0.01

Test interpretation:
H0: The difference between the proportions is equal to 0.
Ha: The difference between the proportions is lower than 0.
As the computed p-value is greater than the significance level alpha=0.01,
one cannot reject the null hypothesis H0.
The risk to reject the null hypothesis H0 while it is true is 9.27%.

We compare these values to the ones shown in the text. We note that these values are exactly the same as the values shown on the MINITAB printout.

8.5 Technology Lab

The Technology Lab consists of problems for the student to practice the techniques presented in each lesson. Each problem is taken from the homework exercises within the *Statistics* text and includes an **Excel** data set (when applicable) that should be used to create the desired output. The completed output has been included with each problem so that the student can verify that he/she is generating the correct output.

1. **Cooling method for gas turbine.** During periods of high demand for electricity – especially in the hot summer months – the power output from a gas turbine engine can drop dramatically. One way to counter this drop in power is by cooling the inlet air to the turbine. AN increasingly popular cooling method uses high-pressure inlet fogging. The performance of a sample of 67 gas turbines augmented with high-pressure inlet fogging was investigated in the *Journal of Engineering for Gas Turbines and Power* (Jan. 2005). One measure of performance is heat rate (kilo-joules per kilowatt hour). Heat rates for the 67 gas turbines, saved in the GASTURBINE file, are listed below. Suppose that a standard gas turbine has, on average, a heat rate of 10,000 kJ/kWh. Conduct a test to determine whether the mean heat rate of gas turbines augmented with high-pressure inlet fogging exceeds 10,000kJ/kWh. Use $\alpha = .05$.

XLSTAT Output

One-sample t-test / Upper-tailed test:

Difference	1066.4328
t (Observed value)	5.4729
t (Critical value)	1.6683
DF	66
p-value (one-tailed)	< 0.0001
alpha	0.05

Test interpretation:
H0: The difference between the means is equal to 0.
Ha: The difference between the means is greater than 0.
As the computed p-value is lower than the significance level alpha=0.05,
one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.
The risk to reject the null hypothesis H0 while it is true is lower than 0.01%.

2. **Dissolved organic compound in lakes**. The level of dissolved oxygen in the surface water of a lake is vital to maintaining the lake's ecosystem. Environmentalists from the University of Wisconsin monitored the dissolved oxygen levels over time for a sample of 25 lakes in the state (*Aquatic Biology*, May 2010). To ensure a representative sample, the environmentalists focused on several lake characteristics, including dissolved organic compound (DOC). The DOC data (measured in grams per cubic-meters) for the 25 lakes are listed in the table and found in the WISCLAKES data file. The population of Wisconsin lakes has a mean DOC value of 15 grams/m^3. Use a test of hypothesis to make an inference about whether the sample is representative of all Wisconsin lakes for the characteristic, dissolved organic compound. Use $\alpha = .05$.

LAKE	DOC
Allequash	9.6
BigMuskellunge	4.5
Brown	13.2
Crampton	4.1
CranberryBog	22.6
Crystal	2.7
EastLong	14.7
Helmet	3.5
Hiawatha	13.6
Hummingbird	19.8
Kickapoo	14.3
LittleArborVitae	56.9
Mary	25.1
Muskellunge	18.4
NorthgateBog	2.7
Paul	4.2
Peter	30.2
Plum	10.3
ReddingtonBog	17.6
Sparkling	2.4
Tenderfoot	17.3
TroutBog	38.8
TroutLake	3
Ward	5.8
WestLong	7.6

XLSTAT Output

One-sample t-test / Two-tailed test:

95% confidence interval on the mean:
(9.1649, 19.8671)

Difference	-0.4840
t (Observed value)	-0.1867
|t| (Critical value)	2.0639
DF	24
p-value (Two-tailed)	0.8535
alpha	0.05

Test interpretation:
H0: The difference between the means is equal to 0.
Ha: The difference between the means is different from 0.
As the computed p-value is greater than the significance level alpha=0.05, one cannot reject the null hypothesis H0.
The risk to reject the null hypothesis H0 while it is true is 85.35%.

3. **Satellite radio in cars**. A spokesperson for the National Association of Broadcasters (NAB) claims that 80% of all satellite radio subscribers have a satellite radio receiver in their car. A June, 2007 survey of 501 satellite radio subscribers found that 396 have a satellite receiver in their car. Test to determine if the true percentage of all satellite radio subscribers that have a satellite radio receiver in their car is less than 80%. Use $\alpha = .10$.

XLSTAT Output

z-test for one proportion / Lower-tailed test:

Difference	-0.0096
z (Observed value)	-0.5361
z (Critical value)	-1.2816
p-value (one-tailed)	0.2959
alpha	0.1

Test interpretation:
H0: The difference between the proportions is equal to 0.
Ha: The difference between the proportions is lower than 0.
As the computed p-value is greater than the significance level alpha=0.1,
one cannot reject the null hypothesis H0.
The risk to reject the null hypothesis H0 while it is true is 29.59%.