

LAB 4 INSTRUCTIONS

CONFIDENCE INTERVALS & HYPOTHESIS TESTING

In the lab instructions, we will explore how to create confidence intervals and perform hypothesis tests using statistical tools in Excel to make inferences for one and two-sample problems.

For **Activating the Data Analysis Add-In** or **Inserting Excel Output into a Word Document**, see the **Lab 1 instructions**.

1. One-Sample Confidence Intervals in Excel

The purpose of a confidence interval is to estimate an unknown population parameter with an indication of how accurate the estimate is and of how confident we are the result is correct. The confidence interval has the following form.

$$\text{estimator} \quad \pm \quad \underbrace{(\text{critical value}) \times (\text{standard error})}_{\text{margin of error}}$$

The critical value is determined from the confidence level. The confidence level states the success rate of the *method* used to construct a confidence interval containing the population parameter. That is, if you use 95% confidence intervals often, in the long run, 95% of your intervals will contain the true parameter value. You cannot know whether a particular confidence interval contains the parameter. The estimator and standard error are typically determined from the sample data. The margin of error is the product of the critical value and standard error.

The standard error can sometimes be based off parameter values, such as the population standard deviation (σ); if it is known, the z -distribution is used to determine the critical value. More realistically, however, the standard error is estimated from the sample, using both the sample standard deviation (s) and the sample size (n); in these situations, the t -distribution is used to determine the critical value.

$$\sigma \text{ known} \\ \bar{X} \pm z_{\alpha/2} \left(\frac{\sigma}{\sqrt{n}} \right)$$

$$\sigma \text{ unknown} \\ \bar{X} \pm t_{n-1, \alpha/2} \left(\frac{s}{\sqrt{n}} \right)$$

1.1 One-Sample Confidence Intervals using Descriptive Statistics

The **Descriptive Statistics** feature (already discussed at length in the **Lab 1 Instructions**) can directly calculate the margin of error for a confidence interval, using the sample standard deviation and the t -distribution. Checking *Confidence Level for Mean* and adding the desired confidence level (see below) will add one more row to the output.

Below is the exact same output as seen in the **Lab 1 Instructions** with the addition of the margin of error for a two-sided 95% confidence interval. Thus, the two-sided 95% confidence interval can be displayed as 68.92 ± 4.812662 ; quick calculations produce the lower and upper bounds of the interval to re-display the interval as (64.10734, 73.73266).

<i>Scores</i>	
Mean	68.92
Standard Error	2.394865
Median	72.5
Mode	75
Standard Deviation	16.93425
Sample Variance	286.769
Kurtosis	-0.21749
Skewness	-0.60693
Range	71
Minimum	28
Maximum	99
Sum	3446
Count	50
Confidence Level(95.0%)	4.812662

Note: When inputting a confidence level, Excel assumes the corresponding confidence interval will be two-sided. If creating a one-sided confidence interval/bound, careful consideration will be needed. For example, if a one-sided 95% confidence interval/bound is needed, then the desired confidence level should be 90%.

1.2 One-Sample Confidence Intervals using Excel functions

First review the **Lab 1 Instructions** to see how to respectively use the AVERAGE, STDEV, and COUNT functions for the sample mean, sample standard deviation, and sample size.

Note that if constructing a confidence interval where σ is known, the sample standard deviation is not required in the calculations; rather, a population standard deviation value is needed, likely stated somewhere in the question. Also, this confidence interval permits the use of the NORM.S.INV or NORMSINV functions. Please review the **Lab 3 Instructions** to see how to use them fully, yet a brief example is provided here for guidance.

For a 95% two-sided confidence interval, $1 - \alpha = 0.95$ and $z_{\alpha/2} = z_{0.025}$. In other words, $P(Z < z) = 0.025$.

Thus, $z = \text{NORM.S.INV}(0.025) = -1.95996$ or, alternatively, $z = \text{NORM.S.INV}(1 - 0.025) = 1.95996$.

As the confidence interval is two-sided, either critical value calculation will work since they have the exact same magnitude despite different signs. If the confidence interval is one-sided, carefully consider what sign would apply.

Using the *Scores* output above, $68.92 \pm (1.95996) \times (2.394865) = 68.92 \pm 4.693850 = (64.22615, 73.61385)$. This interval does not equal the one above since this uses the z -distribution, not the t -distribution. The margin of error is notably lower (as expected) due to theory relating these two distributions.

If constructing a confidence interval where σ is unknown, then the T.INV or TINV functions can be used. Either function takes two arguments as described below or via Microsoft ([T.INV](#) and [TINV](#)). Note that the two functions do not work exactly the same.

T.INV(probability,deg_freedom)

The T.INV function syntax has the following arguments:

probability The probability associated with the Student's t-distribution.
deg_freedom The number of degrees of freedom with which to characterize the distribution.

The arguments in the T.INV function must satisfy the following conditions: **probability** is a number between 0 and 1 while **deg_freedom** is a number greater than or equal to one. If **deg_freedom** is not an integer, it will be truncated.

Example:

Using the *Scores* output above, construct a 90% two-sided confidence interval.

For a 95% two-sided confidence interval, $1 - \alpha = 0.95$ and $t_{\alpha/2, n-1} = t_{0.025, 49}$.

Thus, $t = \text{T.INV}(0.025, 49) = -2.009575$ or, alternatively, $t = \text{T.INV}(1 - 0.025, 49) = 2.009575$.

Using the *Scores* output above, $68.92 \pm (2.009575) \times (2.394865) = 68.92 \pm 4.812662 = (64.10734, 73.73266)$. This interval is exactly the same as the one produced by the **Descriptive Statistics** feature because the exact same margin of error is produced from both methods.

`TINV(probability,degrees_freedom)`

The TINV function syntax has the following arguments:

probability	The probability associated with the two-tailed Student's t-distribution.
degrees_freedom	The number of degrees of freedom with which to characterize the distribution.

The arguments in the TINV function must satisfy the following conditions: **probability** is a number between 0 and 1 while **deg_freedom** is a number greater than or equal to one. If **deg_freedom** is not an integer, it will be truncated.

Example:

Using the *Scores* output above, construct a 90% two-sided confidence interval.

For a 95% two-sided confidence interval, $1 - \alpha = 0.95$ and $\alpha = 0.05$. Also, $df = n - 1 = 50 - 1 = 49$.

Thus, $t = \text{TINV}(0.05, 49) = 2.009575$.

Using the *Scores* output above, $68.92 \pm (2.009575) \times (2.394865) = 68.92 \pm 4.812662 = (64.10734, 73.73266)$. This interval is exactly the same as the one produced by the **Descriptive Statistics** feature because the exact same margin of error is produced from both methods.

2. One-Sample Hypothesis Tests in Excel

In any hypothesis testing problem there are two alternative claims under consideration: the null hypothesis denoted by H_0 (the claim initially assumed to be true) and the alternative hypothesis denoted by H_A (the claim we suspect is true instead of H_0). Both hypotheses are statements about a population parameter; our goal is to determine which of them is more consistent with the sample data. Two possible actions regarding H_0 : reject or not reject H_0 .

The following errors can be made: A Type I error consists of rejecting the null hypothesis H_0 when it is true. A Type II error involves not rejecting H_0 when H_0 is false. The probability of a Type I error is also called the significance level and is denoted by α . Given a fixed value of α , we try to find tests producing high probability that the test will reject H_0 when, in fact, H_0 is false. This probability is called the power of the test.

As with the confidence interval, there are situations where σ is known, yet since these situations are less likely, the focus will be on situations where σ is unknown; thus, the t -distribution will be used. Consider a normal population with an unknown population mean and unknown population standard deviation. In order to test $H_0: \mu = \mu_0$ based on a simple random sample size (n), the test statistic (t_0) is defined as follows.

$$t_0 = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

Thus, the value of t_0 is the distance between the sample mean and the hypothesized value in terms of the number of standard deviations. The larger the absolute value of t_0 , the stronger the evidence against the null hypothesis and the smaller the p -value. The p -value is a measure of plausibility between the hypothesized value for a population characteristic and the observed sample. The smaller the p -value, the more likely we reject H_0 .

The test statistic (t_0) follows a t -distribution with $n - 1$ degrees of freedom. There is no feature in Excel that allows calculating the value of the test statistic automatically. You have to enter the above formula into Excel worksheet to obtain the value of the test statistic for a given sample size (n). First obtain the value of the sample mean and sample standard deviation for the sample (see earlier in the **Lab 4 Instructions**). The statistic is based on the sample mean, which becomes approximately normal as the sample size gets larger even when the population does not have a normal distribution.

To calculate the p -value of the hypothesis test (to measure the strength of the evidence against the null hypothesis), the TDIST or T.DIST functions can be used. Either function takes three arguments as described below or via Microsoft ([TDIST](#) and [T.DIST](#)). Note that the two functions do not work the same and TDIST is preferred.

TDIST(x,degrees_freedom,tails)

The TDIST function syntax has the following arguments:

x	The numeric value at which to evaluate the distribution.
degrees_freedom	The number of degrees of freedom with which to characterize the distribution.
tails	The number of tails in the distribution.

The arguments in the TDIST function must satisfy the following conditions: **x** is a number greater than 0 (use the ABS function whether or not the value is negative, see below) while **deg_freedom** is a number greater than or equal to one. If **deg_freedom** is not an integer, it will be truncated. The values of **tails** is either 1 (one-tailed) or 2 (two-tailed).

Example:

Using the *Scores* output above, test the following two-tailed hypothesis of $H_0: \mu = 68.5$ vs. $H_A: \mu \neq 68.5$.

$$t_0 = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{68.92 - 68.5}{16.93425 / \sqrt{50}} = \frac{0.42}{2.394865} = 0.175375$$

Thus, p -value = TDIST(ABS(0.175375), 50 - 1, 2) = 0.861508. The p -value is fairly high (close to 1) so regardless of the approach used in the full hypothesis test, H_0 would not be rejected.

T.DIST(x,deg_freedom, cumulative)

The T.DIST function syntax has the following arguments:

x	The numeric value at which to evaluate the distribution.
degrees_freedom	The number of degrees of freedom with which to characterize the distribution.
cumulative	A logical value that determines the form of the function. If cumulative is TRUE, T.DIST returns the cumulative distribution function; if FALSE, it returns the probability density function.

The arguments in the T.DIST function must satisfy the following conditions: **x** is a number greater than 0 (use the ABS function whether or not the value is negative, see below) while **deg_freedom** is a number greater than or equal to one. If **deg_freedom** is not an integer, it will be truncated. Lastly, **cumulative** is either TRUE or FALSE.

Example:

Using the *Scores* output above, test the following two-tailed hypothesis of $H_0: \mu = 68.5$ vs. $H_A: \mu \neq 68.5$.

$$t_0 = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{68.92 - 68.5}{16.93425 / \sqrt{50}} = \frac{0.42}{2.394865} = 0.175375$$

Thus, p -value = 2*(1 - T.DIST(ABS(0.175375), 50 - 1, 1)) = 2*(1 - 0.569246) = 0.861508. The p -value is fairly high (close to 1) so regardless of the approach used in the full hypothesis test, H_0 would not be rejected.

3. Two-Sample Hypothesis Tests in Excel

Consider two normal populations with unknown means μ_1 and μ_2 . The **Data Analysis** menu in Excel includes the following four hypothesis testing procedures to test $H_0: \mu_1 - \mu_2 = \delta_0$, where δ_0 is the hypothesized difference.

- (a) t-Test: Paired Two Sample for Means
- (b) t-Test: Two-Sample Assuming Equal Variances
- (c) t-Test: Two-Sample Assuming Unequal Variances
- (d) z-Test: Two Sample for Means

In order to use any of the four testing procedures, first enter data into two columns in a spreadsheet, and then enter their ranges into a dialog box. Note that option (d) requires knowing the population variances, which is not realistic, so this option will not be reviewed.

3.1 Two Independent Sample Hypothesis Tests in Excel

Excel offers a choice between two t -tests for independent samples (options (b) and (c)). One is labeled for equal variances, the other for unequal variances. The unequal variances test is the two-sample t -test in which the population variances are unknown and not assumed equal. The test is valid whether or not the population variances are equal, yet the name remains misleading as it suggests “assuming unequal variances” rather than “not assuming equal variances”.

The other test is a special version of the two-sample t -test that assumes that the two population variances are equal. The test is based on a more accurate (pooled) estimate of the common standard deviation and produces slightly narrower confidence intervals and slightly more powerful tests. As the sample sizes get bigger, the advantages make less and less difference.

The two-sample t -test with a pooled standard deviation is slightly more powerful than the two-sample t -test without equal variances, but serious error can result if the standard deviations are not equal. You can “assume equal variances” and use the **t-Test: Two-Sample Assuming Equal Variances** option if all of the three following conditions hold. If at least one of the conditions does NOT hold, the non-pooled t -test is more appropriate.

- (1) The sample sizes are approximately equal (within 1-2 observations),
- (2) Both sample sizes are at least 15,
- (3) The ratio of the larger and smaller standard deviations is less than two ($s_{\max} / s_{\min} < 2$).

As in the **Lab 1 Instructions**, suppose the first 25 values of *Scores* came from Section A and the last 25 values came from Section B. Note that the sample sizes are equal, both are at least 15, and the standard deviation ratio produces a value less than 2 ($19.99558/12.77693 \approx 1.564795$). Thus, it would be possible to assume equal variances.

Using the **t-Test: Two-Sample Assuming Equal Variances** feature within the **Data Analysis** feature and clicking on OK will present a dialog box exactly the same or similar to the following one. The **Variable 1 Range** should be the range presented by the cells containing *Scores* for Section A. The **Variable 2 Range** should be the range presented by the cells containing *Scores* for Section B. The ‘up’ arrow to the right of the box can be clicked to permit the mouse to select the data directly from the spreadsheet. The **Hypothesized Mean Difference** is typically a value of 0. The **Labels** box should be checked if labels are present. For the **Output Options**, it is preferred to select *Output Range* and choose a single cell to present the upper-left corner of all the corresponding output. Then click **OK**. To adjust the width of a column to fit the longest entry, double-click the column heading border between the column and the next column.

For comparison, using the **t-Test: Two-Sample Assuming Unequal Variances** feature within the **Data Analysis** feature and clicking on OK will present a dialog box exactly the same or similar to the following one. As seen, it's essentially the same as the dialog box above.

Both outputs are produced below, yet note that the one on the left is more appropriate. Both outputs provide the sample means, sample variances, and sample sizes of each group. The “pooled procedure” also provides a pooled variance while the two also differ in their values for df due to different calculations. The equal sample sizes (because the standard error and difference between the means would be equal) permit the test statistics to be equal, yet the different df identify different one-tailed p -values as well as two-tailed p -values. Depending on the approach to make a conclusion regarding the hypotheses, the numbers for the current p -values may result in a variety of conclusions. Consult course notes to proceed further.

t-Test: Two-Sample Assuming Equal Variances

	A	B
Mean	65.64	72.2
Variance	399.8233	163.25
Observations	25	25
Pooled Variance	281.5367	
Hypothesized Mean Difference	0	
df	48	
t Stat	-1.38227	
P(T<=t) one-tail	0.086644	
t Critical one-tail	1.677224	
P(T<=t) two-tail	0.173289	
t Critical two-tail	2.010635	

t-Test: Two-Sample Assuming Unequal Variances


	A	B
Mean	65.64	72.2
Variance	399.8233	163.25
Observations	25	25
Hypothesized Mean Difference	0	
df	41	
t Stat	-1.38227	
P(T<=t) one-tail	0.087187	
t Critical one-tail	1.682878	
P(T<=t) two-tail	0.174375	
t Critical two-tail	2.019541	


3.2 Paired Sample Hypothesis Tests in Excel

The paired samples test is used when the same objects or subjects are observed twice. A common application is to take measurements before and after some kind of treatment. Using the **t-Test: Paired Two Sample for Means** feature within the **Data Analysis** feature and clicking on OK will present a dialog box exactly the same or similar to the following one. The **Variable 1 Range** should be the range presented by the cells containing the first half of the paired samples. The **Variable 2 Range** should be the range presented by the cells containing the second half of the paired samples. The ‘up’ arrow to the right of the box can be clicked to permit the mouse to select the data directly from the spreadsheet. The **Hypothesized Mean Difference** is typically a value of 0. The **Labels** box should be checked if labels are present. For the **Output Options**, it is preferred to select **Output Range** and choose a single cell to present the upper-left corner of all the corresponding output. Then click **OK**.

t-Test: Paired Two Sample for Means

Input

Variable 1 Range: 


Variable 2 Range: 

Hypothesized Mean Difference:

☐ Labels

Alpha:

Output options

☒ Output Range: 

☐ New Worksheet Ply:

☐ New Workbook

OK Cancel Help

It is also possible to obtain a confidence interval of the differences by creating a new column of differences from the two existing columns of data. Since a single column of differences can be considered a single sample, return to the **One-Sample Confidence Intervals in Excel** section above to use either way of finding the confidence interval.