CHAPTER 17 Appendix

Inference about a Population Mean

In this chapter, we remove the (usually unrealistic) assumption that the population standard deviation σ is known. This requires a new distribution: the t distribution. Not only are we estimating the population mean μ from the data, but we are also estimating the population standard deviation. In our discussions, the t distributions are all centered about 0; their variability is governed by a new parameter, the *degrees of freedom* (df). For one sample, df = $n - 1$.

As always, there are assumptions. The main one here is that the sample mean $\overline{x}$ has (at least approximately) a Normal distribution. This can be accomplished with small sample sizes by having an (approximately) Normal distribution in the population, or through the central limit theorem by having a large sample size. You should always check a plot of the data with small sample sizes: Is the Normal population assumption reasonable? Check that your data are approximately symmetric with no strong outliers. You can use a Normal plot (discussed in Chapter 11), a histogram, a dotplot, or boxplot (discussed in Chapters 1 and 2).

t Confidence Interval for Mean

Excel

Confidence intervals are available only with the Analysis Toolpak add-in, and **only with raw data**. Excel does not compute the interval as such, but does calculate the margin of error.

1. **Data ➔ Data Analysis ➔ Descriptive Statistics**
2. In the Input Range section, enter or select the range of data.
3. Check the boxes for **Summary Statistics** and **Confidence Level for Mean**. Enter the desired confidence level.
4. **OK**. Excel returns the margin of error, labeled "Confidence Level (xx.x%)." Add and subtract this from the sample mean to form the confidence interval.

***t* Test for Population Mean**

Excel does not provide a dedicated statistical function or routine for testing the mean when the population standard deviation is unknown.

Matched Pairs *t* Test

1. **Data ➔ Data Analysis ➔ t-Test: Paired Two Sample for Means**
2. Select or enter the range of cells for the first variable—the first observation from each pair.
3. Select or enter the range of cells for the second variable—the second observation from each pair.
4. Enter a hypothesized mean difference, usually 0.
5. **OK**. Excel returns the mean and variance of each sample, the degrees of freedom (df), t statistic, and two P-values (both one- and two-tailed) as well as the two associated critical values.

For more help and examples, see the Excel Video Technology Manuals on *One Sample Mean Inference t—summarized data, One Sample Mean Inference t—with data*, and *Paired Samples Inference*.

With raw data (entered in a column):
You can use either method described here with raw data. One typically would use Analyze Distribution, as that platform is more powerful.

***t* Confidence Interval for Mean**

1. **Analyze ➔ Distribution**
2. Click to select a column, then **Y, columns ➔ OK**.
3. The 95% confidence interval is shown in the summary statistics section as "Lower and Upper 95% Mean."
4. To change the confidence level, select the Red Triangle next to the variable name ➔ **Confidence Interval**. This allows a confidence level of 90%, 95%, and 99%, or you can click **Other** and enter a different confidence level.

***t* Test for Population Mean**

1. **Analyze ➔ Distribution**
2. Click to select a column, then **Y, columns ➔ OK**.
3. Select the Red Triangle, and click **Test Mean**.
4. Enter the hypothesized value of the mean.
5. **OK**. In addition to the sample mean and standard deviation, JMP returns the degrees of freedom, *t* statistic, and three P-values: one for each one-tail test and one for the two-tail test. Select the correct *P*-value for your test.

Matched Pairs *t* Test

1. **Analyze ➔ Matched Pairs**
2. Select the two numeric columns with paired data, then click **Y, Paired Response** to enter both variables in the *y* role.
3. **OK**. JMP returns a plot of the differences, both sample means, the ends of the confidence interval, the *t*-statistic, and three *P*-values as described previously.

With summarized data (You are given $\bar{x}$, *s*, and *n*).

***t* Confidence Interval for Mean**

1. **Help ➔ Calculators ➔ Confidence Interval for One Mean**
2. Move the radio button to **Summary Statistics** and click **OK**.
3. Make sure the radio button at left is on ***t***; if not, click to move it.
4. Enter the sample average ($\bar{x}$), sample standard deviation (s), and the sample size (n).
5. The default confidence level is 95% (0.95). If you want a different level, change that parameter. Press the Tab or Enter key to update the results.

Hypothesis Test for One Mean

6. **Help ➔ Calculators ➔ Hypothesis Test for One Mean**
7. Move the radio button to **Summary Statistics** and click **OK**.
8. Make sure the radio button at left is on ***t***; if not, click to move it. Move the second radio button to correspond to the correct form of the alternate hypothesis.
9. Enter the hypothesized mean μ_o, sample average ($\bar{x}$), sample standard deviation (s), and the sample size (n), and significance level (α).
10. Press the Tab or Enter key to update the results. If you check the box at left to Reveal Decision, JMP will give you the Reject/Fail to Reject the null hypothesis decision in red below the numeric test results.

For more help and examples, see the JMP Video Technology Manuals on *One Mean Test—t test—summarized data, One Mean Test—t test—with data, One Sample Mean Inference t—with data, One Sample Mean Inference t—summarized data*, and *Paired Samples Inference*.

Minitab

t Confidence Interval and Tests for One Sample Mean

Confidence intervals and tests are performed from the same dialog box regardless of whether you have raw data (in a column) or only summary statistics.

1. **Stat ➔ Basic Statistics ➔ 1-Sample t**
2. Change the drop-down at the top to reflect whether you have summary statistics or raw data in a column.
3. For raw data, highlight and **Select** the column name into the box.
4. For summary statistics, enter the sample size (n), sample mean ($\bar{x}$), and standard deviation (s).
5. The defaults are 95% confidence and a "not equal" alternative (two-sided confidence interval). To change either of these, click **Options**, change the value, and then click **OK**.
6. For the *confidence interval*, click **OK**.
7. For a *hypothesis test*, check the box, enter the value of μ_0 and then click **OK**.

Matched Pairs *t* Test and Confidence Interval

Output here includes the test statistic for Group 1 – Group 2, as well as the confidence interval for the mean difference. **Paired data must be in two separate columns**. If you already have a column of differences, use the methods described previously for a single sample with a hypothesized difference of 0.

1. **Stat ➔ Basic Statistics ➔ Paired t**
2. Change the drop-down at the top to reflect whether you have summary statistics or raw data in two columns.
3. For raw data, highlight and **Select** each column name into a box (Minitab will form differences as Sample 1 – Sample 2),
4. For summary statistics, enter the sample size (n), sample mean of the differences ($\bar{x}$), and standard deviation of the differences (s).
5. The defaults are 95% confidence, a hypothesized difference of 0, and a "not equal" alternative (two-sided confidence interval). To change either of these, click **Options**, change the value, and then click **OK**.
6. Click **OK.** Minitab returns summary statistics for each sample as well as for the differences, along with the confidence interval and test results.

Note: With raw data, clicking **Graphs** in the main dialog box will allow you to produce graphs of the data to check assumptions such as Normality.

For more help and examples, see the Minitab Video Technology Manuals *One Sample Mean Inference t—with data, One Sample Mean Inference t—summarized data*, and *Paired Samples Inference*.

SPSS

Confidence Interval for Mean

1. **Analyze ➔ Descriptive Statistics ➔ Explore**
2. Select the column of interest, and press the arrow to enter the column into the **Dependent List**.
3. The default is 95% confidence. To change that, click **Statistics** and enter a new confidence level. **Continue**.
4. **OK**

Test for Population Mean

1. **Analyze ➔ Compare Means ➔ One Sample T-Test**
2. Select the column of interest, and press the arrow to enter the column into the **Test Variable(s)** section.
3. Enter the hypothesized mean in the **Test Value** box.
4. **OK**. The "Confidence Interval for the Difference" in the output is for the difference $\mu - \mu_0$. SPSS gives only a two-tailed *P*-value (labeled "Sig"). Assuming the sample mean is on the "correct" side of the hypothesized value, divide that by 2 for a one-tailed test.

Matched Pairs *t* Test and Confidence Interval

Output here includes the test statistic for Group 1 – Group 2, as well as the confidence interval for the mean difference. **Paired data must be in two separate columns**.

1. **Analyze ➔ Compare Means ➔ Paired-Samples T-Test**
2. Select the two columns of interest (paired responses), and press the arrow to enter the columns into the **Paired Variable** section. If you want to change the confidence level, click **Options** and enter the desired value.
3. **OK.** SPSS gives only a two-tailed *P*-value (labeled "Sig"). Divide that by 2 for a one-tailed test.

For more help and examples, see the SPSS Video Technology Manuals *One Sample Mean Inference t—with data* and *Paired Samples Inference*.

CRUNCHIT!

Both confidence intervals and hypothesis tests are done through the **Statistics ➔ t ➔ 1-Sample** dialog.

***t* Confidence Interval and Tests for One Sample Mean**

1. For "raw data" in a column, click the **Columns** tab; otherwise, click **Summarized**.
2. For raw data, select the column of interest.
3. For summarized data, enter the sample size (n), observed sample mean, and sample standard deviation.
4. Select the **Confidence Interval** tab and enter the desired confidence level.
5. Alternatively, select the **Hypothesis Test** tab. Enter the hypothesized mean value and type of alternative hypothesis.
6. **Calculate**

Matched Pairs *t* Test

1. **Statistics ➔ t ➔ Paired**
2. Select the first and second variables.
3. Select the **Hypothesis Test** tab or the **Confidence Interval** tab.
4. For a test, enter the mean difference under the null hypothesis (usually 0) and type of alternative hypothesis. For an interval, enter the desired confidence level.
5. **Calculate**

For more help and examples, see the CrunchIt! Help Videos *Confidence Intervals Using t, One Mean Test Using t*, and *Paired Samples Inference*.

TI-83/-84

Confidence Interval for a Mean

1. Press STAT ▶ ▶ to select **TEST**.
2. Select **8:TInterval**.
3. Select **Data** (you have entered the actual sample data in a list) or **Stats**.
4. For *raw data*, enter L1 (or whichever list has your data) as the data location, and the confidence level.
5. For *summarized data* (you have Stats), enter the sample mean ($\bar{x}$), the sample standard deviation (s), the sample size (n), and the confidence level.
6. **Calculate**

Test for Population a Mean

1. Press STAT ▶ ▶ to select **TEST**.
2. Select **2:T-test**.
3. Select **Data** (you have entered the actual sample data in a list) or **Stats**.
4. Enter the value of the mean under the null hypothesis.
5. For *raw data*, enter L1 (or whichever list has your data) as the data location.
6. For *summarized data* (you have Stats), enter the sample mean, sample standard deviation, and sample size.
7. Select the desired type of alternative hypothesis.
8. **Calculate**

Matched Pairs *t* Test

1. Enter data in L1 and L2.
2. Move to the top of L3, and enter **L1-L2**; this will calculate the difference for each row.
3. Proceed as described earlier, using the column of differences (L3) as your data.

For more help and examples, see the TI Video Technology Manuals *One Sample Mean Inference t—with data, One Sample Mean Inference t—summarized data*, and *Paired Samples Inference*.

***t* Confidence Interval and Tests for One Sample Mean**

Both confidence intervals and hypothesis tests are done through the t.test command.

```
> t.test(var, alternative='two.sided', mu=15,
conf.level=.95,data=dataframe)
```

where "var" is the name of the variable that is part of the "dataframe" data set; possible alternatives are "two.sided," "less," or "greater." If the "mu = " portion is omitted, a test against a mean of 0 is computed by default, but this allows computation of just the confidence interval.

Matched Pairs t Test

Both confidence intervals and hypothesis tests are done through the t.test command. Note the addition of "paired=T" in the command.

```
> t.test(x,y,alternative="two.sided",paired=T,conf.
level=0.95,data=dataframe)
```

For more help and examples, see the R Video Technology Manuals *Inference for a Population Mean based on t Using Raw Data, Significance Test for a Population Mean Based on t Using Summarized Data*, and *Paired Samples Inference*.