

# Hypothesis Testing Workbook

## Exercise 1.1 – Testing Claims about a Mean

1. Using Minitab®, work along with the instructor to conduct the 1-sample z test for the “Lead Time” example:

When ordering commodities, the use of a corporate credit card has traditionally been assumed to generate the best response. That assumption is now going to be tested against a standard requisition form process.

Historically, when utilizing the corporate credit card:

Average delivery time = 6 days

Standard deviation = 2 days

A random sample of size 36 was collected from the requisition process.

The data is in column C1 of [Hypothesis Testing & Minitab – Data Files.MPJ](#)

Using Minitab®...

Select: **Stat>Basic Statistics>1-sample z**

Under “**Samples in Columns**”, insert **C1:Delivery**

Under “**Standard Deviation**”, enter **:2** (known from historical data)

Under “**Test mean**”, enter **:6** (known from historical data)

Select “**Graphs**” and check Histogram of data, click **OK**

Select “**Options**” and note the defaults:

Confidence level is 95%

Alternative hypothesis is “**not equal**”.

Both of these apply to our current example

Click **OK**, click **OK** again

Is the response time for the requisition form process the same as the response time when using a corporate credit card?

If the processes are different, which is better (ie. faster)?

## Exercise 1.1 (continued)

2. Break into small teams. Conduct the appropriate hypothesis test for each of these examples:

### **Example A: Cell Phone Charges**

(Data is in column C2 of [Hypothesis Testing & Minitab – Data Files.MPJ](#))

Cell phone and pager costs have escalated in recent years. A procedural change occurred in how cell phones are issued and a study was conducted to determine if the average bill has changed from the previous level of \$75.00 per month with a standard deviation of \$15.00. A random sample of charges for 50 phones was conducted and the following results were obtained:

$$\bar{X} = \$76.09 \quad s = 9.57$$

Is the process average, as it is running today, different than the previous level at a 5% level of significance?

What is the appropriate hypothesis test for this case?

- a) Establish the Alternative and Null Hypotheses:

H<sub>0</sub>:

H<sub>1</sub>:

- b) Level of significance,  $\alpha = 0.05$  (given)
- c) Randomly select a representative sample of data. This was done. Data is in column C2 of [Hypothesis Testing & Minitab – Data Files.MPJ](#).
- d) Compute the P-value: the probability of obtaining the observed sample if the null hypothesis is true.

Using Minitab®...

Select: **Stat>Basic Statistics>1-sample z**

Under “**Sample in Columns**”, insert C2: **Phones (\$)**

- What value should be entered under “Standard Deviation”? \_\_\_\_\_
- What value should be entered under “Test mean”? \_\_\_\_\_
- Under “Options”, are the defaults applicable for this case? \_\_\_\_\_

## Exercise 1.1 (continued)

- e) Compare the P-value to the level of significance,  $\alpha$

---

---

Is the new process average different from the previous level?

---

---

### **Example B: Steel Bracket Rolling Mill**

(Data is in column C3 of BB-HypTest.MTW)

For a steel bracket rolling mill, there is a target inside dimension of 0.75 inches. Today, you go and sample 25 brackets and obtain the following results:

$$\bar{X} = 0.738 \quad s = 0.019$$

Is the process average, as it is running today, different than the target at a 1% level of significance?

What is the appropriate hypothesis test for this case?

- a) Establish the Alternative and Null Hypotheses:

$H_0$ :

$H_1$ :

- b) Determine the level of significance,  $\alpha$

$\alpha = 0.01$  (given)

- c) Randomly select a representative sample of data. This was done. Data is in C3 of [Hypothesis Testing & Minitab – Data Files.MPJ](#). Note: Since this is a small sample, run a test for normality.

- d) Compute the P-value: the probability of obtaining the observed sample if the null hypothesis is true.

Using Minitab®...

Select: **Stat>Basic Statistics>1-sample t**

Under "**Sample in Columns**", insert **C3: Dimension**

- What value should be entered for the test mean? \_\_\_\_\_
- Under "Options", are the defaults applicable in this case? \_\_\_\_\_

- e) Compare the P-Value to  $\alpha$ . Is the process running on target? \_\_\_\_\_

## Exercise 1.2 – Estimating Population Means and Differences Between Population Means

### 1: Pivot Torque (Estimating Population Mean, Small Sample)

You are testing the pivot torque for a Boat Propellar. You want to confirm that the average torque is greater than 30 Newton-meters. Is it?

From a random sample of 16 screws (data is in Column C4 of [Hypothesis Testing & Minitab – Data Files.MPJ](#)):

$$\begin{aligned}\bar{X} &= 32.32 \text{ Newton-meters} \\ s &= 3.25 \text{ Newton-meters}\end{aligned}$$

Using Minitab®...

1. Conduct Anderson-Darling test for normality.
2. Provide a point estimate for the population mean:

$$\mu = \bar{x} = \underline{\hspace{2cm}}$$

3. Find confidence intervals at:

90% Confidence ( $\alpha = 0.10$ )

95% Confidence ( $\alpha = 0.05$ )

99% Confidence ( $\alpha = 0.01$ )

Minitab® commands (data is in column C4 of [Hypothesis Testing & Minitab – Data Files.MPJ](#)):

### **Select: Stat > Basic Statistics > 1–Sample t**

Under “Samples in columns:” enter column containing data. (C4)  
See “Options...” to set confidence level (95% is the default).

## Exercise 1.2 – Estimating Population Means and Differences Between Population Means (Cont.)

Example A – Total Processing Time for Purchase Orders (in days)

It is the department's goal to completely process all purchase orders within 30 days. You track the total processing time for a sample of 100 purchase orders and obtain the following results:

Average of the sample =  $\bar{x} = 26.9$  days

Sample standard deviation =  $s = 4.9$  days

Based on this sample, what is the estimated average processing time for all purchase orders? Provide a 95% confidence interval for the estimate.

Point estimate of the population mean:  $\mu = \bar{x} =$  \_\_\_\_\_

Margin of Error:  $E = Z_{\alpha/2} \left[ \frac{s}{\sqrt{n}} \right] =$

\_\_\_\_\_

95% Confidence Interval: \_\_\_\_\_  $< \mu <$  \_\_\_\_\_

Minitab® Commands (data is found in column C5 of [Hypothesis Testing & Minitab – Data Files.MPJ](#)):

Conduct Normality Test

Select: **Calc > Column Statistics > Standard Deviation**  
to calculate the standard deviation

Then select: **Stat > Basic Statistics > 1–Sample Z**

Under “**Samples in columns:**” enter column containing data

For “**Standard deviation:**”, enter the computed standard deviation from above.

Note: The default is 95% confidence level (See **Options...**)

Click **OK**

## Example B: Milled pipe diameter

For a given pipe, there is a target inside dimension of 0.75 inches. You sample 49 pipes from the rolling mill and obtain the following results:

Average of the sample =  $\bar{y} = 0.741$  inches

Sample standard deviation =  $s = 0.017$  inches

Is the process average, as it is running today, different from the target at the 95% confidence level?

Point estimate of population mean:  $\mu = \bar{x} =$  \_\_\_\_\_

Margin of Error:  $E = Z_{\alpha/2} \left[ \frac{s}{\sqrt{n}} \right] =$  \_\_\_\_\_

95% Confidence Interval: \_\_\_\_\_  $< \mu <$  \_\_\_\_\_

Minitab® Commands (Data is found in column C6 of [Hypothesis Testing & Minitab – Data Files.MPJ](#)):

Conduct Normality Test

Select: **Calc > Column Statistics > Standard Deviation**  
to calculate the standard deviation

Then select: **Stat > Basic Statistics > 1–Sample Z**

Under “**Samples in columns:**” enter column containing data

For “**Standard deviation:**”, enter the computed standard deviation from above.

Note: The default is 95% confidence level (See **Options**)

Click **OK**

## Exercise 1.3: Two Population Means

### Example A: Processing Time (from the training module: work together)

A Black Belt conducted a study to determine if training was effective at reducing the time it takes to process orders. The process is not automated and is highly dependent on the knowledge and skill level of individual processors.

The orders used throughout the study are essentially identical in magnitude. Data represents processing time, in hours.

<u>Processor (C9)</u>	<u>Before Training (C10)</u>	<u>After Training (C11)</u>
1	23	21
2	17	11
3	8	6
4	9	10
5	7	5
6	25	22
7	16	11
8	11	16
9	12	10
10	9	5

Are the two samples dependent (paired data) or independent (unpaired data)?

What is the appropriate hypothesis test in this case?

a) Establish the Alternative and Null Hypotheses.

$H_0$ :

$H_1$ :

b) Determine the level of significance,  $\alpha$

Use  $\alpha = 0.05$ .

c) Randomly select representative samples. This was done. Data is in columns C10 and C11 of [Hypothesis Testing & Minitab – Data Files.MPJ](#).

Since the sample size is small ( $n_1$  and  $n_2$  are both  $<30$ ), run a test for normality on each sample.

## Exercise 1.3 (continued)

- d) Compute the P-value: the probability of obtaining the observed sample if the null hypothesis is true.

Using Minitab®...

Select: **Stat>Basic Statistics>Paired t**

Under "**First sample**", insert **C10: Before Training**

Under "**Second sample**", insert **C11: After Training**

View "**Options**" and make sure the proper confidence level and alternative

hypothesis are set.

Select: "**Graphs**" and check Histogram of differences.

- e) Compare the P-value to the level of significance,  $\alpha$ .
- 
- 

Does the data provide sufficient evidence that the training results in improved (ie. reduced) processing time?

## Exercise 1.3 (continued)

### Example B: Molded Component

A Black Belt is studying the outputs of an injection molding operation where the critical dimension was measured and we were using two different colors depending on the ship type being supplied.

Red	Gray
46.80	49.61
48.88	49.28
49.82	50.31
52.32	51.09
46.97	52.70
48.33	49.01
48.44	48.95
50.34	52.03
50.82	51.24
52.14	51.84

The complete samples are found in Columns C12 “red” and C13 “gray” of [Hypothesis Testing & Minitab – Data Files.MPJ](#) ( $n_1 = n_2 = 200$ ).

Suppose we wanted to test the hypothesis that the data come from different populations. We must set up a test that will demonstrate a statistically significant difference between the two sample means.

Use a level of significance  $\alpha = 0.05$  for this test.

What is the appropriate hypothesis test in this case?

a) Establish the Alternative and Null Hypotheses:

$H_0$ :

$H_1$ :

b) Determine the level of significance,  $\alpha$ .

$\alpha =$

c) Randomly select a representative sample of data. This was done. The data is in columns C12 and C13 of [Hypothesis Testing & Minitab – Data Files.MPJ](#). A sample size of  $n_1 = n_2 = 200$  was used.

d) Compute the P-Value: the probability of obtaining the observed sample if the null hypothesis is true.

Using Minitab®...

Select: **Stat>Basic Statistics>2-Sample t**

View “**Options**” and make sure the proper values are set.

P-Value = \_\_\_\_\_

## Exercise 1.3 (continued)

e) Compare the P-Value to the level of significance,  $\alpha$ .

---

---

Did the data come from two different populations? \_\_\_\_\_

### **Example C: Shipping Cost**

A Black Belt conducted a study to compare the average monthly dollar figure per part for shipping before and after procedural changes to materials handling.

Part Number	Monthly Shipping Cost	
	Old Procedures	New Procedures
1	4050	3550
2	10280	10100
3	14700	13500
4	8380	8300
5	4300	4150
6	6790	6750
7	10500	9050
8	14800	12080
9	13930	13070
10	11050	9580

Are the two samples dependent (paired data) or independent (unpaired data)?

What is the appropriate hypothesis test in this case?

a) Establish the Alternative and Null Hypotheses.

$H_0$ :

$H_1$ :

b) Determine the level of significance,  $\alpha$

Use  $\alpha = 0.05$ .

## Exercise 1.3 (continued)

- c) Randomly select a representative sample. This was done. Data is in columns C14 and C15 of [Hypothesis Testing & Minitab – Data Files.MPJ](#).

Since the sample size is small ( $n_1$  and  $n_2$  are both  $<30$ ), run a test for normality on each sample.

- d) Compute the P-value: the probability of obtaining the observed sample if the null hypothesis is true.

Using Minitab®...

Select: **Stat>Basic Statistics>Paired t**

Under “**First sample**”, insert **C14: Old**

Under “**Second sample**”, insert **C15: New**

View “**Options**” and make sure the proper confidence level and alternative

hypothesis are set.

Select: “**Graphs**” and check Histogram of differences.

- e) Compare the P-value to the level of significance,  $\alpha$ .
- 
- 

Does the data provide sufficient evidence that the new procedures result in lower shipping costs?

## Exercise 1.3 (continued)

### Example C: Shipping Cost – Teaching Point

The previous example, shipping cost, involved dependent (paired) samples. The paired t test is the appropriate hypothesis test in such a case.

Using the paired t test, you obtained a P-value smaller than the significance level  $\alpha = 0.05$ . Therefore, you concluded that the new procedures in materials handling resulted in lower shipping cost.

What if you incorrectly ran a 2-sample t test (intended for independent, unpaired, samples) instead of the paired t test? Would you be led to draw a different conclusion?

For illustration purposes, go ahead and run the 2-sample t test with this data. But remember, THIS IS THE WRONG TEST FOR THIS APPLICATION!

Can you explain why the 2-sample t test does not allow us to see the difference in cost between the old and new procedures?

---

---

---

---

## Exercise 1.4 – Population Means and Variances

1. Break into small teams.
2. For each example provided, determine the appropriate hypothesis test, conduct the test and draw conclusions.

### **Example A: Cure Time of Rubber**

Four rubber recipes are being tested for cure time. All have been determined to have the appropriate density to be considered for use in new propulsor gaskets. Use Columns C19 “f1”, C20 “f2”, C21 “f3”, and C22 “f4” of [Hypothesis Testing & Minitab – Data Files.MPJ](#) to determine if there is a significant difference in average cure time for any of the formulations at the 5% level of significance.

Are the samples independent or dependent?

What is the appropriate hypothesis test in this case?

- a) Establish the Alternative and Null Hypotheses.  
H<sub>0</sub>:  
H<sub>1</sub>:
- b) Determine the level of significance,  $\alpha$ .  
 $\alpha = 0.05$  (given)
- c) Randomly select a representative sample of data. This was done. Data is in Columns C19, C20, C21, and C22 of [Hypothesis Testing & Minitab – Data Files.MPJ](#).

Compute the P-value: the probability of obtaining the observed sample if the null hypothesis is true.

The Minitab® one-way ANOVA feature requires the data to be stacked in one column, so be sure to use one-way (unstacked).

Select: **Stat>ANOVA>One-way (unstacked)**

Under “**Response**”, insert Columns C19, C20, C21, and C22.

Click **OK**

## Exercise 1.4 (Continued)

- d) Compare the P-value to the level of significance,  $\alpha$ .

---

---

Does the data provide sufficient evidence that a difference in average cure time exists among the four rubber formulations?

---

- e) If a significant difference is detected, then conduct Tukey's test of pairwise comparisons to see which formulations differ from the others.

Select: **Stat>ANOVA>One-way (unstacked)**  
Insert the same columns as before for "**Response**".  
Select "**Comparison**".  
Check  $\sqrt{}$ Tukey's, family error rate.  
Click **OK**, Click **OK** again.

Look for the Lower and Upper ranges that do NOT contain zero. These are the relationships that are significantly different.

Which foam formulations differ from one another?

---

---

---

### **Example B: Oil Density**

A Black Belt is studying two different oil formulations to be used in a new buoyancy system. She is interested in determining which oil formulation yields a more consistent density. Thus, she needs to compare the population variances.

Two samples were randomly selected across multiple lots for each oil type. The data come from normal distributions.

Use the Hypothesis Testing Decision Trees to determine the appropriate hypothesis test for this case.

## Exercise 1.4 (Continued)

- a) Establish the Alternative and Null Hypotheses.  
H<sub>0</sub>:  
H<sub>1</sub>:
- b) Determine the level of significance,  $\alpha$ .  
Use  $\alpha = 0.05$
- c) Randomly select a representative sample of data. This was done. Data is in columns C23 (OilA) and C24 (OilB) of [Hypothesis Testing & Minitab – Data Files.MPJ](#).

Run the Anderson-Darling test for normality on both columns, to verify the two populations are normally distributed.

- d) Compute the P-value: the probability of obtaining the observed sample if the null hypothesis is true.

Using Minitab®...

Select: **Stat>Basic Statistics>2 Variances**

Select: **“Samples in different columns”**.

Under **“First”**, insert column C23: **OilA**

Under **“Second”**, insert Column C24: **OilB**

View **“Options”** and make sure desired Confidence Level is set.

Click **OK**, Click **OK** again.

Notice that Minitab® gives two P-values, one for the F test and one for Levene’s test. Which is the more valid test in this case?

- e) Compare the P-value to the significance level,  $\alpha$ .

---

---

---

Does the data provide sufficient evidence that there is a difference in variance between the two oil formulations?

---

Which formulation has a more consistent density?

---

## Exercise 1.5 - Testing Equality of Population Proportions

1. Break into small teams
2. For the examples provided, perform the appropriate hypothesis tests.

### Example : Processing Cell Phone and Pager Requests

A Black Belt is comparing two methods of processing cell phone and Blackberry requests to determine if there is a difference in methods..

Method	No. Audited	Number with Errors	Proportion Defective
Standard Paper	80	12	0.150
Intranet Form	80	5	0.063

Is there a significant difference between the methods at the 5% level of significance?

- What type of data is being used? \_\_\_\_\_
  - What is the appropriate hypothesis test? \_\_\_\_\_
- a) Establish the Alternative and Null Hypotheses.  
H<sub>0</sub>:  
H<sub>1</sub>:
  - b) Determine the level of significance,  $\alpha$ .  
 $\alpha = 0.05$  (given)
  - c) Randomly select a representative sample of data. This was done. Data is shown in above table.
  - d) Compute the P-value: the probability of obtaining the observed sample if the null hypothesis is true.

Using Minitab®...

Select: **Stat>Basic Statistics>2 proportions**

Select: **"Summarized data"**

Under **"First Sample"**, enter **80 "Trials"** and **12 "Events"**

## Exercise 1.5 (Continued)

Under "**Second Sample**", enter **80** "**Trials**" and **5** "**Events**"

View "**Options**" and make sure the proper settings are selected. –

Pooled p

Click **OK**

- e) Compare the P-value to the level of significance,  $\alpha$ .

---

---

Does the data provide sufficient evidence that one method is more accurate than the other? \_\_\_\_\_

## Exercise 1.6 – Testing Equality of Population Proportions

### Example : Scrap Rate

A Black Belt is studying the scrap rate across multiple shifts. He wants to determine if there is a statistically significant difference among the three shifts at the 5% level of significance.

Shift	Parts Produced	Number Defective	Proportion Defective
1	3240	128	0.0395
2	3100	114	0.0368
3	3460	102	0.0295

- What type of data is being used?
  - Use the Hypothesis Testing Decision Trees to determine the appropriate test.
- a) Establish both the Alternative and Null Hypotheses.  
H<sub>0</sub>:  
H<sub>1</sub>:
- b) Determine the level of significance,  $\alpha$ .  
 $\alpha = 0.05$  (given)
- c) Randomly select a representative sample of data. This was done. Data is shown in above table. It is also contained in columns C25, C26 and C27 of [Hypothesis Testing & Minitab – Data Files.MPJ](#)
- d) Construct a Chi-squared test.

Using Minitab®...

Select: **Stat>Tables>Chi-Squared Test**

Under columns containing data, insert , C26, C27

Click **OK**.

P-value is compared to alpha 0.05. IF > then no evidence of difference between shifts

- e) What is your conclusion?
-

## Exercise 1.7 – Testing for Association

### Example A: Critical Blueprint Defects

1. Your Value Stream Champion is concerned that work on the new DDX is resulting in more critical blueprint defects than previous programs. The Program Manager disagrees. You decide to resolve the argument once and for all. Based on their input, you select two similar programs and review print revisions to evaluate four different kinds of critical print defects, then compare the results to the DDX prints. Results are listed below. Next, you conduct a Chi-Square test for Critical Blueprint Defects. (Data is in columns C28, C29, C30 and C31 of [Hypothesis Testing & Minitab – Data Files.MPJ](#)).

Shift	Defect A	Defect B	Defect C	Defect D	Total
Pgrm A	15	21	45	13	94
Pgrm B	26	31	34	5	96
DDX	33	17	49	20	119
<i>Total</i>	74	69	128	38	309

Does the type of defect that occurs depend on the Program?

H<sub>0</sub>: the defect type is independent of the Program.

H<sub>1</sub>: Program and defect type are dependent.

Using Minitab®...

Select: **Stat>Tables>Chi-Square Test**

Under “**Columns containing the table:**”, insert **C28: Defect A, C29: Defect B, C30: Defect C and C31: Defect D**

Click **OK**

Is there a dependency between defect type and program?

## Exercise 1.7 (Continued)

### Example B: NMCI Defects

Over the course of the past six months, 309 defects were found on NMCI seats. They took on four different types (A,B,C,and D). We wish to investigate whether the defects vary from location to location. The counts are summarized in the table below. (Note: the total seats over this time period were virtually identical by location.)

Location	Defect A	Defect B	Defect C	Defect D	Total
1	15	21	45	13	94
2	26	31	34	5	96
3	33	17	49	20	119
Total	74	69	128	38	309

Does the type of defect that occurs depend on the shift?

$H_0$  : the defect is independent of location

$H_1$  : location and defect type are dependent

Select: **Stat > Tables > Chi-square test**

Select Columns C44, C45, C46, & C47 and click **OK**.

Do the defects vary by location?

## More Practice:

### Order Filling Process (Estimating the Difference Between Two Population Means, Large Samples)

Use Minitab® to work through the following examples:

You are comparing two approaches for processing orders: the existing (old) approach and a new, automated approach.

Is there a difference between the two approaches in terms of average processing time?

You have collected data (in hours) on 35 orders using the old approach and 35 orders using the new approach. (Data is in columns C50 & C51 of [Hypothesis Testing & Minitab – Data Files.MPJ](#)).

Using Minitab®...

1. Estimate the difference between the two population means (old approach minus new approach).
2. Provide a 95% confidence interval around the estimate of the difference. Does the new approach seem to save time?

Minitab® commands (data is in columns C50 & C51 of [Hypothesis Testing & Minitab – Data Files.MPJ](#)):

No need to do Normality – Avg. of Avg. – forces normality

**Select: Stat > Basic Statistics > 2-Sample t**

(NOTE: Minitab® does not have a 2-Sample Z test)

**Select “Samples in Different Columns”**

- Under “First”, enter one of the columns of data.
- Under “Second”, enter the other column of data.

See “Options” to set confidence level (95% is the default). Click **OK**. Check the p-value. Is there a difference between the two approaches?

## More Practice (Continued)

### Program Management (Estimating the Difference Between Two Population Means, Small Samples)

You are comparing Value Streams in terms of the time Green Belts actually spend on Lean Six Sigma activities. You have a sample of 8 Green Belts from Value Stream A and 8 Green Belts from Value Stream B.

Data is “Percentage of time spent on Lean Six Sigma activities” for each Green Belt (and can be found in columns 54 and 55 of [Hypothesis Testing & Minitab – Data Files.MPJ](#)).

Using Minitab®...

1. Estimate the difference between the two population means.
2. Provide a 95% confidence interval around the estimate of the difference. Does there appear to be a difference, on average, between the time spent on Green Belt activities for the two Value Streams?

Minitab® Commands (data is in columns C54 and C55 of [Hypothesis Testing & Minitab – Data Files.MPJ](#)).

**Select: Stat > Basic Statistics > 2-Sample t...**

**Select “Samples in Different Columns”**

- Under “First”, enter one of the columns of data.
- Under “Second” enter the other columns of data.

See “Options” to set confidence level (95% is the default). Click **OK**. Check the p-value. Is there a difference between the two approaches?