

# Hypothesis Testing and Confidence Intervals

## Example 1: Infusion Amount

### Problem

A medical device company has developed a new IV with an automatic pump that can be programmed to deliver a fixed amount of antibiotic fluid. A particular device is programmed to deliver an amount of 365 mL/hr. Determine if the actual fluid delivered by this device is significantly different from this set target amount.

### Data collection

An IV with an automatic pump is set for an infusion amount of 365 mL/hr and run for one hour. The amount of fluid passing through the IV per hour is measured for six different 1-hour periods.

### Tools

- **Display Descriptive Statistics**
- **1-Sample t**
- **Normality Test**
- **Help**
- **Individual Value Plot**

### Data set

InfusionAmountvsTarget.MPJ

Variable	Description
Infusion Amount	The amount of fluid delivered by the IV over 1 hour (mL)

## Displaying descriptive statistics

Use descriptive statistics to summarize important features of the data. In particular, descriptive statistics provide useful information about the location and variability of the data.

### Display Descriptive Statistics

1. Open InfusionAmountvsTarget.MPJ.
2. Choose **Stat** > **Basic Statistics** > **Display Descriptive Statistics**.
3. In **Variables**, enter '*Infusion Amount*'.
4. Click **OK**.

## Interpreting your results

The statistics indicate that the sample mean is 366.70 mL/hr. This is slightly above the target amount of 365 mL/hr. The one-sample t-test will compare this difference (1.7 mL/hr) to the variation in the data. While the standard deviation of the data is 2.40 mL/hr, the one-sample t-test will use the SE Mean.

### Statistics

Variable	N	N*	Mean	SE Mean	StDev	Minimum
Infusion Amount	6	0	366.70	0.981	2.40	363.43

Variable	Q1	Median	Q3	Maximum
Infusion Amount	364.25	367.22	368.31	370.13

## Standard Error of the Mean

We can see that the SE Mean is simply the standard deviation divided by the square root of the number of data points and represents the dispersion or variation in the distribution of sample means. The one-sample t-test uses the distribution of the sample mean (not the distribution of the data) for the analysis. Therefore, the standard error of the mean will be used as the estimate of variation for the t-test and confidence interval.

### What is the standard error of the mean?

The standard error of the mean (SE of the mean) estimates the variability between sample means that you would obtain if you took multiple samples from the same population. The standard error of the mean estimates the variability between samples whereas the standard deviation measures the variability within a single sample.

For example, you have a mean delivery time of 3.80 days with a standard deviation of 1.43 days based on a random sample of 312 delivery times. These numbers yield a standard error of the mean of 0.08 days (1.43 divided by the square root of 312). Had you taken multiple random samples of the same size and from the same population the standard deviation of those different sample means would be around 0.08 days.

Use the standard error of the mean to determine how precisely the mean of the sample estimates the population mean. Lower values of the standard error of the mean indicate more precise estimates of the population mean. Usually, a larger standard deviation will result in a larger standard error of the mean and a less precise estimate. A larger sample size will result in a smaller standard error of the mean and a more precise estimate.

Minitab uses the standard error of the mean to calculate the confidence interval, which is a range of values likely to include the population mean.

# Hypothesis testing

## What is a hypothesis test

A hypothesis test uses sample data to test a hypothesis about the population from which the sample was taken. The 1-sample t-test is one of many procedures available for hypothesis testing in Minitab.

For example, to test whether the mean length of stents is equal to the target length, measure several stents and use the mean length of these samples to estimate the mean length of the total stent population. Using information from a sample to make a conclusion about a population is known as statistical inference.

## When to use a hypothesis test

Use a hypothesis test to make inferences about one or more populations when sample data are available.

## Why use a hypothesis test

A hypothesis test answers questions such as:

- Is a process centered correctly?
- Is the product from one supplier better than the product from another?
- Do treatment groups in an experiment differ?

For example,

- Is the mean width of a medical device the desired 8.5 inches?
- Do seals from one supplier hold longer than seals from another supplier?
- Do customers prefer one formulation of a children's antibiotic flavor enhancer over another?

# 1-sample t-test

## What is a 1-sample t-test

Use a 1-sample t-test to determine whether  $\mu$  (the population mean) is equal to the hypothesized mean.

The test uses the standard deviation of the sample to estimate  $\sigma$  (the population standard deviation). If the difference between the sample mean and the hypothesized mean is large relative to the variability of the sample mean, then  $\mu$  is unlikely to be equal to the hypothesized mean.

## When to use a 1-sample t-test

Use a 1-sample t-test with continuous data from a single random sample.

The test assumes the population is normally distributed. However, the test is robust to violations of this assumption, provided the observations are collected randomly and the data are continuous, unimodal, and reasonably symmetric [1].

## Why use a 1-sample t-test

A 1-sample t-test answers questions such as:

- Is a process on target?
- Does a key characteristic of a supplier's material have the desired mean value?

For example,

- Is the mean width of syringe needles higher or lower than target?
- Is the mean strength of the seal from a supplier lower than the required minimum?

## Testing the null hypothesis

The company needs to determine whether the mean differs significantly from the target of 365. In statistical terms, the process mean is the *population mean*, or  $\mu$  (mu).

### Statistical hypotheses

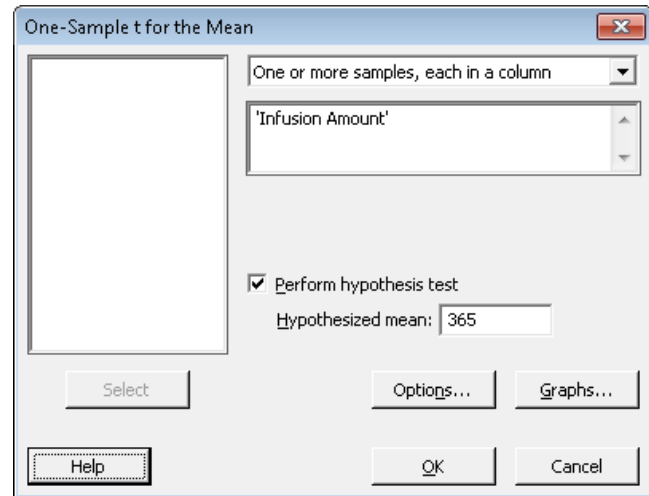
Either  $\mu$  is equal to 365, or it is not. You can state these alternatives as two hypotheses:

- The *null hypothesis* ( $H_0$ ):  $\mu$  is equal to 365.
- The *alternative hypothesis* ( $H_1$ ):  $\mu$  is *not* equal to 365.

Because the company cannot feasibly measure every IV device in the population, they can never know for certain which hypothesis is correct. However, an appropriate hypothesis test can help them make an informed decision. For the infusion data, the appropriate test is a 1-sample t-test.

### 1-Sample t

1. Choose **Stat** > **Basic Statistics** > **1-Sample t**.
2. Complete the dialog box as shown below.



3. Click **OK**.

## Interpreting your results

### The logic of hypothesis testing

All hypothesis tests follow the same steps:

1. Assume  $H_0$  is true.
2. Determine how different the sample is from what you expected under the above assumption.
3. If the sample is sufficiently unlikely under the assumption that  $H_0$  is true, then reject  $H_0$ .

For example, the t-test results indicate that the sample mean is 366.705. The test answers the question, "If  $\mu$  is equal to 365, how likely is it to see a sample mean this different (or even more different) from 365?" The answer is given as a probability value (P), which for this test is equal to 0.143.

### T

The t-statistic (1.74) is:

$$t = (\text{sample mean} - \text{hypothesized mean}) / \text{SE Mean}$$

where SE Mean is the standard error of the mean (a measure of variability). For a given sample size, as the value of t increases, the p-value decreases.

### Descriptive Statistics

N	Mean	StDev	SE Mean	95% CI for $\mu$
6	366.705	2.403	0.981	(364.183, 369.226)

$\mu$ : mean of Infusion Amount

### Test

Null hypothesis  $H_0: \mu = 365$

Alternative hypothesis  $H_1: \mu \neq 365$

T-Value	P-Value
1.74	0.143



## Interpreting your results

### Making a decision

To make a decision, choose the significance level,  $\alpha$  (alpha), before the test:

- If P is less than or equal to  $\alpha$ , reject  $H_0$ .
- If P is greater than  $\alpha$ , fail to reject  $H_0$ . (Technically, you never *accept*  $H_0$ , you simply fail to reject it.)

A typical value for  $\alpha$  is 0.05, but you can choose higher or lower values depending on the sensitivity required for the test and the consequences of incorrectly rejecting the null hypothesis. Assuming an  $\alpha$ -level of 0.05 for the infusion amount data, not enough evidence is available to reject  $H_0$ . P (0.143) is greater than  $\alpha$ .

### What's next

Check the assumption of normality.

### Descriptive Statistics

N	Mean	StDev	SE Mean	95% CI for $\mu$
6	366.705	2.403	0.981	(364.183, 369.226)

$\mu$ : mean of Infusion Amount

### Test

Null hypothesis  $H_0: \mu = 365$   
 Alternative hypothesis  $H_1: \mu \neq 365$

T-Value	P-Value
1.74	0.143

## Testing the assumption of normality

The 1-sample t-test assumes the data are sampled from a normally distributed population.

Use a normality test to determine whether the assumption of normality is valid for these data.

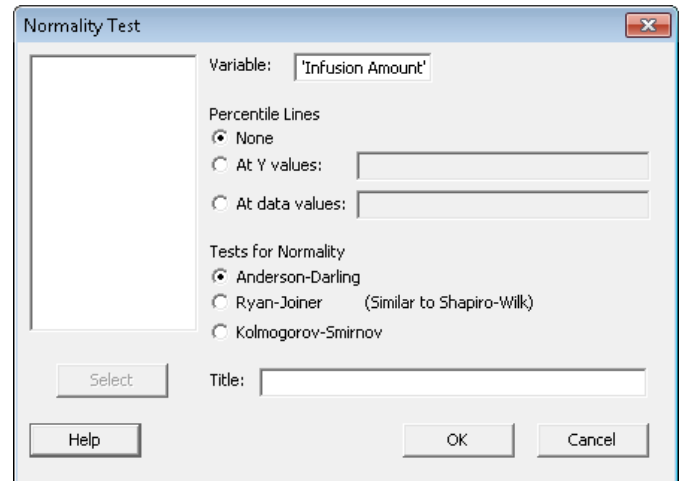
---

**Note** The t-test is robust to violations of the normality assumption, provided that observations are collected randomly and the data are continuous, unimodal, and reasonably symmetric.

---

### Normality Test

1. Choose **Stat > Basic Statistics > Normality Test**.
2. Complete the dialog box as shown below.



3. Click **OK**.

## Interpreting your results

Use the normal probability plot to verify that the data do not deviate substantially from what is expected when sampling from a normal distribution.

- If the data come from a normal distribution, the points will roughly follow the fitted line.
- If the data do not come from a normal distribution, the points will not follow the line.

### Anderson-Darling normality test

The hypotheses for the Anderson-Darling normality test are:

$H_0$ : Data are from a normally distributed population

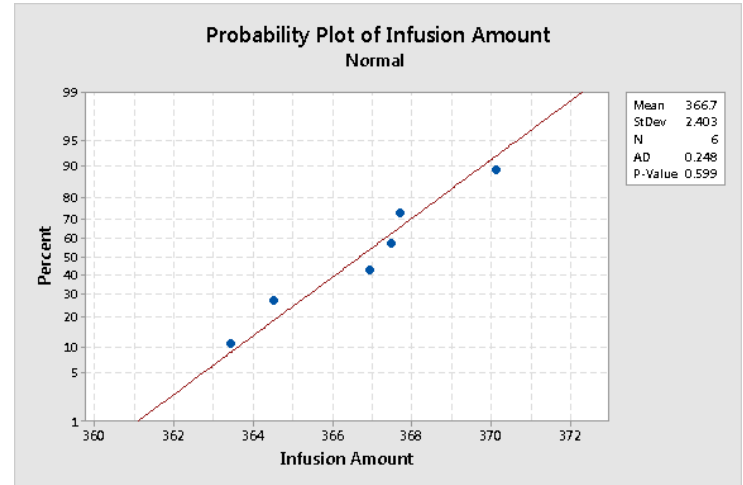
$H_1$ : Data are not from a normally distributed population

Using an  $\alpha$  of 0.05, not enough evidence exists to suggest that the data are not from a normally distributed population.

### Conclusion

Based on the plot and the test, assume that the data are from a normally distributed population.

**Note** When data are not normally distributed, you may be able to transform them using a Box-Cox transformation or use a nonparametric procedure such as the 1-sample sign test.



### What's next

Interpret the confidence interval.

## Confidence intervals

### What is a confidence interval

A confidence interval is a range of likely values for a population parameter (such as  $\mu$ ) that is based on sample data. For example, with a 95% confidence interval for  $\mu$ , you can be 95% confident that the interval contains  $\mu$ . In other words, 95 out of 100 intervals will contain  $\mu$  upon repeated sampling.

### When to use a confidence interval

Use confidence intervals to:

- Make inferences about one or more populations from sample data
- Quantify the precision of the estimate of a population parameter, such as  $\mu$

### Why use a confidence interval

A confidence interval often answers the same questions as a hypothesis test:

- Is  $\mu$  on target?
- How precise is the estimate of  $\mu$ ?
- How low or high might  $\mu$  be?

For example,

- Is the mean tablet weight significantly different from target?
- What is the range of likely values for the mean tablet hardness?

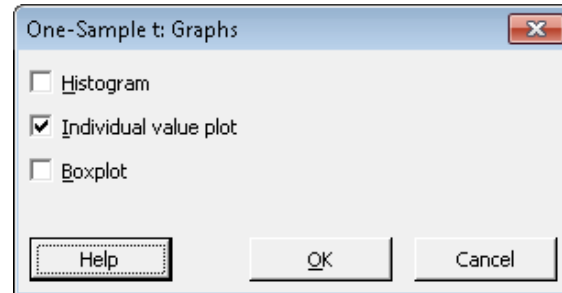
## Using the confidence interval

In the previous analysis, you used a hypothesis test to determine whether the mean was different from the target value. You can also use a confidence interval to evaluate this difference.

The output pane results for 1-Sample t includes values for the upper and lower bounds of the 95% confidence interval. Obtain a graphical representation of the interval by selecting Individual value plot in the Graphs subdialog box.

### 1-Sample t

1. Choose **Stat** > **Basic Statistics** > **1-Sample t**.
2. Click **Graphs**.
3. Complete the dialog box as shown below.



4. Click **OK** in each dialog box.

## Interpreting your results

### Confidence interval

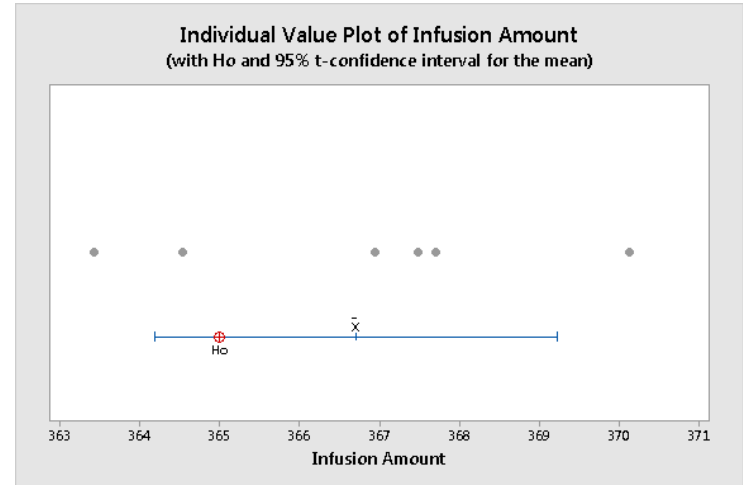
The confidence interval is a range of likely values for  $\mu$ . Minitab displays the interval graphically as a blue line on the individual value plot.

The middle tick mark, labeled  $\bar{X}$ , represents the mean of the sample and the red circle, labeled  $H_0$ , represents the hypothesized population mean (365). Use the confidence interval to test the null hypothesis.

Under repeated sampling from the same population, the confidence intervals from about 95% of the samples would include  $\mu$ . Thus, for any one sample, you can be 95% confident that  $\mu$  is within the confidence interval.

**Note** A confidence interval does not represent 95% of the data; this is a common misconception.

We are 95% confident that the true process mean is between 364.2 mL/hr and 369.2 mL/hr.



## Interpreting your results

### Hypothesis test

- If  $H_0$  is outside the interval, the p-value for the hypothesis test will be less than 0.05. You can reject the null hypothesis at the 0.05  $\alpha$ -level.
- If  $H_0$  is inside the interval, the p-value will be greater than 0.05. You cannot reject the null hypothesis at the 0.05  $\alpha$ -level.

Because the hypothesized population mean, 365, falls within the confidence interval, you cannot reject the null hypothesis. Not enough evidence is available to conclude that  $\mu$  is not 365 at the 0.05 significance level.

### Descriptive Statistics

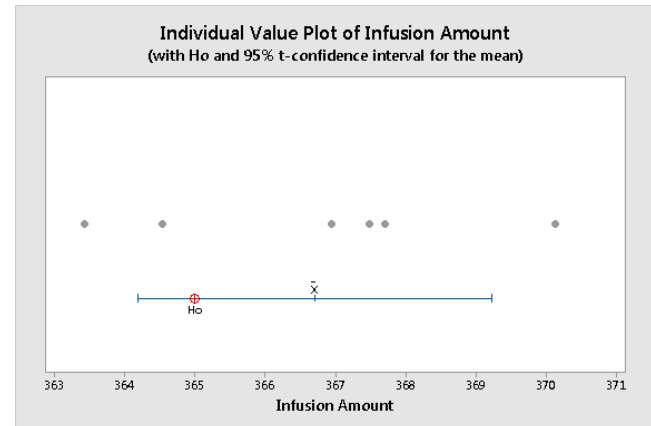
N	Mean	StDev	SE Mean	95% CI for $\mu$
6	366.705	2.403	0.981	(364.183, 369.226)

$\mu$ : mean of Infusion Amount

### Test

Null hypothesis  $H_0: \mu = 365$   
 Alternative hypothesis  $H_1: \mu \neq 365$

T-Value	P-Value
1.74	0.143



## Final considerations

### Summary and conclusions

Based on the sample data, you cannot reject the null hypothesis at the 0.05  $\alpha$ -level. Not enough evidence is available to suggest that the process mean differs from 365.

Like the t-test, the 95% confidence interval does not provide enough evidence to reject the null hypothesis that the population mean is 365. If the process mean is not 365, then it is at least 364.2 and at most 369.2. The team must decide whether these values are reasonable.



## Final considerations

### Hypotheses

A hypothesis test begins with two opposing hypotheses.

The null hypothesis ( $H_0$ ):

- Usually states that some property of a population (such as the mean) is not different from a specified value or from that of another population
- Is assumed to be true until sufficient evidence indicates the contrary
- Is never proven true; you simply fail to disprove it

The alternative hypothesis ( $H_1$ ):

- States that the null hypothesis is wrong
- Can also specify the direction of the difference

### Significance level

Choose the  $\alpha$ -level *before* conducting the test.

- Increasing  $\alpha$  increases the chance of detecting a difference, but it also increases the chance of rejecting  $H_0$  when it is actually true (a Type I error).
- Decreasing  $\alpha$  decreases the chance of making a Type I error, but also decreases the chance of detecting a difference when one actually exists (a Type II error).

### Assumptions

Each hypothesis test is based on one or more assumptions about the data. If these assumptions are not met, the conclusions may not be correct.

When using a 1-sample t-test:

- The sample must be random.
- Sample data must be continuous.
- Sample data are from a normally distributed population.

The t-test is robust to violations of the normality assumption, provided that observations are collected randomly and the data are continuous, unimodal, and reasonably symmetric (see [1]).

### Confidence interval

The confidence interval provides a likely range of values for  $\mu$  (or other population parameters).

You can conduct a 2-tailed hypothesis test (alternative hypothesis of  $\neq$ ) using a confidence interval. For example, if the test value is not within a 95% confidence interval, you can reject  $H_0$  at the 0.05  $\alpha$ -level. Likewise, if you construct a 99% confidence interval and it does not include the test mean, you can reject  $H_0$  at the 0.01  $\alpha$ -level.