One-Sample t-Test

Example 1 Merchandise Shipment Times

Problem

A mail-order catalog company ships orders from a large central warehouse. The warehouse must ensure that the orders are shipped within a specific time frame. If orders reach the shipping bay too quickly, goods may get damaged because of the back-log. Conversely, if orders reach the shipping bay too slowly, they may not be shipped on time, leading to customer dissatisfaction.

Data collection

The warehouse processes orders 24 hours a day. In the course of 3 days, the deviation of the processing time from the target time is recorded for each order.

Deviation = Processing time – Target time

Tools

- **Run Chart**
- **Graphical Summary**
- **1-Sample t**

Data set

CATALOG.MPJ

Variable	Description
Request placed	Date and time the order is placed
Transaction complete	Date and time the order is shipped
Processing time	Time (minutes) taken to process the order (Transaction complete - Request placed)
Target time	Target time (minutes) to ship a particular order
Deviation	Deviation (minutes) of the processing time from the target time

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 one-sample t-test is one of many procedures available for hypothesis testing in MINITAB.

For example, to test whether the mean duration of a transaction is equal to the desired target, measure the duration of several transactions and use the mean of these samples to estimate the mean for all transactions. This is an example of *statistical inference*, which is using information about a sample to make an inference about a population.

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

Hypothesis testing can help answer questions such as:

- Are turn-around times meeting or exceeding customer expectations?
- Is the service at one branch better than the service at another?

For example,

- On average, is a call center meeting the target time to answer customer questions?
- Is the mean billing cycle time shorter at the branch with a new billing process?

One-sample t-test

What is a one-sample t-test

A one-sample t-test helps determine whether μ (the population mean) is equal to a hypothesized value (the test mean).

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

When to use a one-sample t-test

Use a one-sample t-test when continuous data are available from a single random sample.

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

Why use a one-sample t-test

A one-sample t-test can help answer questions such as:

- Is the mean transaction time on target?
- Does customer service meet expectations?

For example,

- On average, is a call center meeting the target time to answer customer questions?
- Is the billing cycle time for a new process shorter than the current cycle time of 20 days?

Testing the randomness assumption

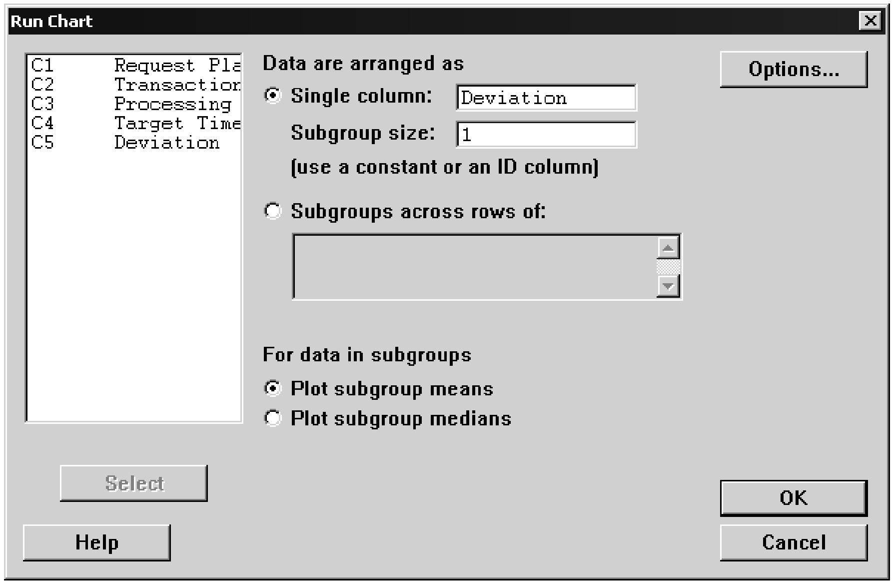
Use a run chart to look for trends or patterns in your data, which may indicate that your data are not random over time.

Run Chart

1 Open CATALOG.MPJ.

2 Choose **Stat ➤ Quality Tools ➤ Run Chart**.

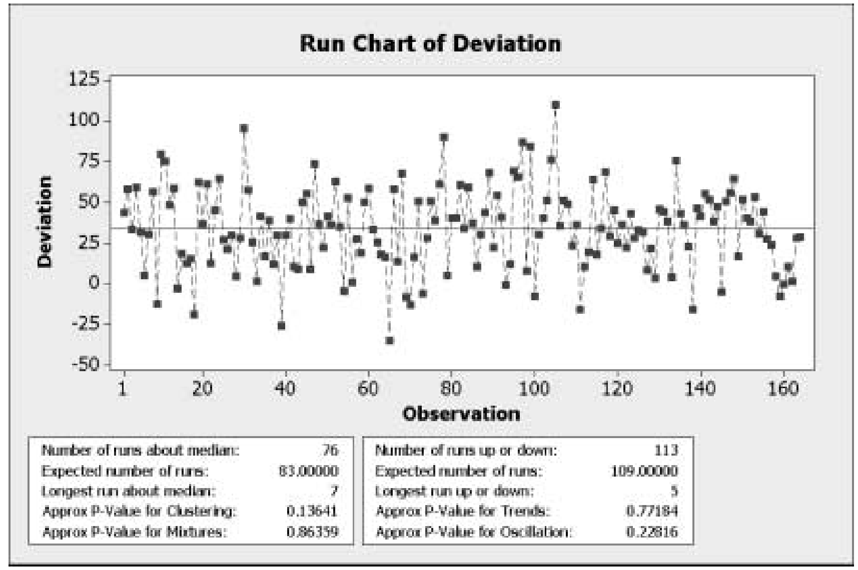
3 Complete the dialog box as shown below.



4 Click **OK**.

Interpreting your results

Run chart performs two tests for randomness that may indicate statistically significant patterns or evidence of nonrandomness.



Test for number of runs about the median

This test is sensitive to two types of nonrandom behavior: mixtures and clusters. Mixtures often indicate combined data from two populations, or two processes operating at different levels. Clusters are groups of points that have similar values.

Test for number of runs up and down

This test is sensitive to two types of nonrandom behavior: oscillation and trend. Oscillation occurs when the data fluctuates up and down rapidly, indicating the process is not steady. Trend is the long-term tendency for a series to rise or fall.

All the p-values associated with the test for nonrandom behavior, are greater than 0.05. You can conclude that no significant nonrandom patterns exist in the data.

Alternate approach

You can also create an I-MR chart with these data. Observation 105, the shipment requested on 3/4/03 at 2:25:49 PM, is above the upper control limit. The shipment is investigated, and no special cause is detected. The anomaly is considered to be a false alarm.

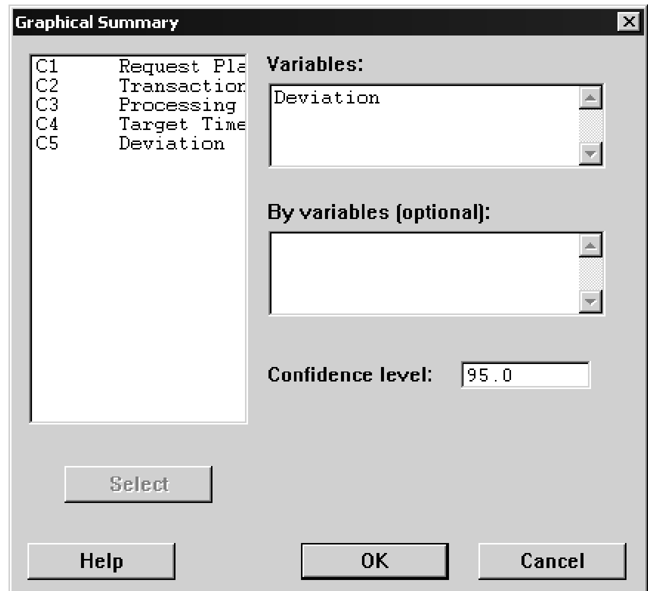
Testing the normality assumption

Use Graphical Summary to verify that the data do not deviate substantially from a normal distribution.

Graphical Summary

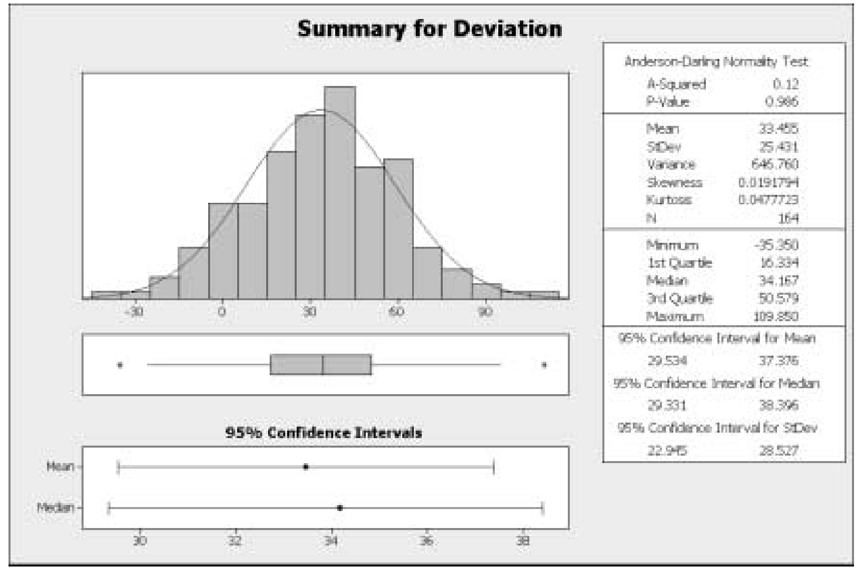
1 Choose **Stat ➤ Basic Statistics ➤ Graphical Summary**.

2 Complete the dialog box as shown below.



3 Click **OK**.

Interpreting your results



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.

The p-value from the Anderson-Darling test (0.986) assesses the probability that the data are from a normally distributed population. Using an α of 0.05, evidence suggests that the data are from a normal population.

Note When the data are not normally distributed, you can transform the data using Box-Cox transformation or use another procedure such as nonparametric tests (1-sample sign test).

Conclusion

Based on the test, you can assume that the data do not deviate substantially from a normal distribution.

What's next

Proceed with the t-test.

Conducting the one-sample t-test

Two-tailed vs. one-tailed tests

The first step in setting up a hypothesis test of means is to determine the null and alternative hypotheses. The null hypothesis usually specifies that a parameter equals a specific value. For example, the average deviation is 0 minutes (H_0: $\mu = 0$ minutes).

The following table shows the three possible alternative hypotheses and the corresponding type of test. You can specify the alternative hypothesis in the Options subdialog box so that MINITAB performs the test correctly.

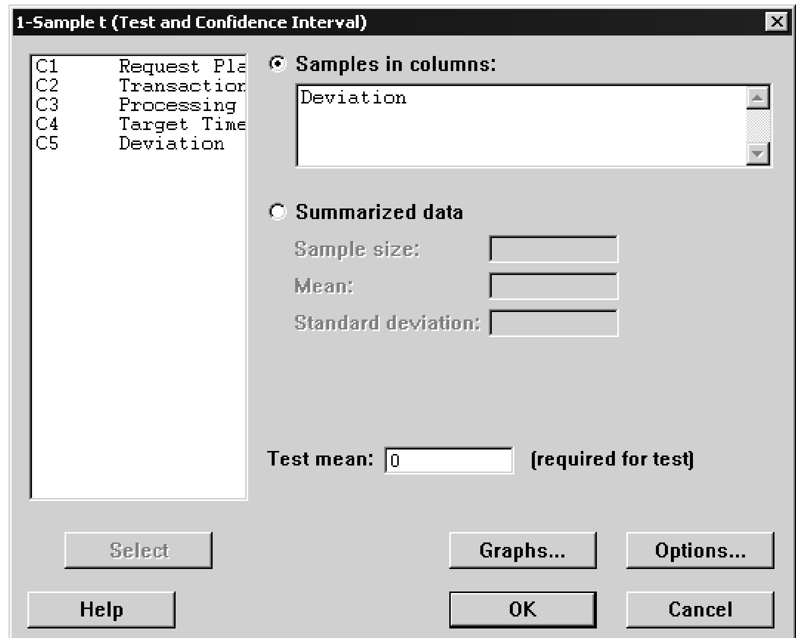
To determine whether the deviation from the target time for shipping is…	Use this test	Alternative hypothesis
Different from 0 minutes	Two-tailed	H_1: $\mu \neq 0$ minutes
Less than 0 minutes	One-tailed (left-tailed)	H_1: $\mu < 0$ minutes
More than 0 minutes	One-tailed (right-tailed)	H_1: $\mu > 0$ minutes

For this example, the two-tailed test answers the question of interest: whether the mean processing time is on target or not. These alternatives can be stated as:

- The *null hypothesis* (H_0): μ is equal to 0 seconds.
- The *alternative hypothesis* (H_1): μ is *not* equal to 0 seconds.

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 testing follows the same steps:

1 Assume H_0 is true.

2 Determine how different the sample is from what is expected under the above assumption.

3 If the sample is sufficiently unlikely under the above assumption, then reject H_0.

Test statistic

The t-statistic (16.85) is calculated as:

t = (sample mean – test mean) / SE Mean

where SE Mean is the standard error of the mean (a measure of variability). As the value of t increases, the p-value becomes smaller.

One-Sample T: Deviation

```
Test of mu = 0 vs not = 0


Variable     N     Mean    StDev  SE Mean        95% CI              T      P
Deviation  164  33.4549  25.4315   1.9859  (29.5335, 37.3762)  16.85  0.000
```

Interpreting your results

Making a decision

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

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

A typical value for α 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.

P-value

The t-test results indicate that the sample mean is not equal to 0 seconds. Thus, the test answers the question, “If μ is equal to 0 seconds, how likely is it to see a sample mean this different (or even more different) from 0 seconds?” The answer is given as a probability value (P), which for this test is equal to 0.000. Because this value is less than α (0.05), you can reject the null hypothesis and conclude that the average deviation is different from 0.

One-Sample T: Deviation

```
Test of mu = 0 vs not = 0


Variable     N     Mean    StDev  SE Mean         95% CI               T      P
Deviation  164  33.4549  25.4315   1.9859  (29.5335, 37.3762)  16.85  0.000
```

Confidence intervals

What is a confidence interval

A confidence interval is a range of likely values for a population parameter (such as μ) that is based on sample data. For example, with a 95% confidence interval for μ, you can be 95% confident that the interval contains μ.

When to use a confidence interval

Use a confidence interval to make inferences about one or more populations from sample data or to quantify the precision of your estimate of μ.

Why use a confidence interval

Confidence intervals can help answer many of the same questions as hypothesis testing:

- Is μ on target?
- How much error exists in an estimate of μ?
- How low or high might μ be?

For example,

- Is the mean transaction time longer than 30 seconds?
- Is the mean daily revenue higher than $6,000?

Interpreting your results

Confidence interval

The 95% confidence interval for the average deviation is between 29.5335 minutes (that is, 29 minutes 32 seconds) and 37.3762 minutes (37 minutes 23 seconds). The 95% confidence interval does not include the target value (zero).

One-Sample T: Deviation

```
Test of mu = 0 vs not = 0


Variable     N     Mean    StDev  SE Mean        95% CI             T      P
Deviation  164  33.4549  25.4315   1.9859  (29.5335, 37.3762)  16.85  0.000
```

Final considerations

Summary and conclusions

Based on the sample data, you can reject the null hypothesis at the 0.05 α-level. The mean of the deviation is significantly different from 0 minutes.

Orders are taking longer than expected, on average. The mean deviation from the target is estimated to be between 29 minutes 32 seconds and 37 minutes 23 seconds, with 95% confidence.

Final considerations

Additional considerations

Hypotheses

A hypothesis test always starts 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 α-level *before* conducting the test.

- Increasing α 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 α decreases the chance of making a Type I error, but also decreases the chance of correctly detecting a difference.

Assumptions

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

When using a one-sample t-test:

- The sample must be random
- Sample data must be continuous
- Sample data should be normally distributed

The t-test procedure is fairly 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 μ (or other population parameters).

You can conduct a two-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 α-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 α-level.