

## Chapter 115

# Confidence Intervals for One Proportion

## Introduction

This routine calculates the sample size necessary to achieve a specified interval width or distance from the sample proportion to the confidence limit at a stated confidence level for a confidence interval for one proportion.

Caution: This procedure assumes that the proportion of the future sample will be the same as the proportion that is specified. If the sample proportion is different from the one specified when running this procedure, the interval width may be narrower or wider than specified.

## Technical Details

Many methods have been devised for computing confidence intervals for a single proportion. Five of these methods are available in this procedure. The five confidence interval methods are

1. Exact (Clopper-Pearson)
2. Score (Wilson)
3. Score with continuity correction
4. Simple Asymptotic
5. Simple Asymptotic with continuity correction

For a comparison of methods, see Newcombe (1998a).

## Confidence Interval Formulas

For each of the following methods, let  $p$  be the population proportion, and let  $r$  represent the number of successes from a sample of size  $n$ . Let  $\hat{p} = r/n$ .

### Exact (Clopper-Pearson)

Using a mathematical relationship (see Fleiss et al (2003), p. 25) between the  $F$  distribution and the cumulative binomial distribution, the lower and upper confidence limits of a  $100(1-\alpha)\%$  exact confidence interval for the true proportion  $p$  are given by

$$\left[ \frac{r}{r + (n - r + 1)F_{1-\alpha/2; 2(n-r+1), 2r}}, \frac{(r + 1)F_{1-\alpha/2; 2(r+1), 2(n-r)}}{(n - r) + (r + 1)F_{1-\alpha/2; 2(r+1), 2(n-r)}} \right]$$

One-sided limits may be obtained by replacing  $\alpha/2$  by  $\alpha$ .

## Confidence Intervals for One Proportion

### Score (Wilson)

The Wilson Score confidence interval, which is based on inverting the z-test for a single proportion, is calculated using

$$\frac{(2n\hat{p} + z_{1-\alpha/2}^2) \pm z_{1-\alpha/2} \sqrt{z_{1-\alpha/2}^2 + 4n\hat{p}(1-\hat{p})}}{2(n + z_{1-\alpha/2}^2)}$$

One-sided limits may be obtained by replacing  $\alpha/2$  by  $\alpha$ .

### Score with Continuity Correction

The Score confidence interval with continuity correction is based on inverting the z-test for a single proportion with continuity correction. The  $100(1-\alpha)\%$  limits are calculated by

$$\text{Lower Limit} = \frac{(2n\hat{p} + z_{1-\alpha/2}^2 - 1) - z_{1-\alpha/2} \sqrt{z_{1-\alpha/2}^2 - \{2 + (1/n)\} + 4\hat{p}\{n(1-\hat{p}) + 1\}}}{2(n + z_{1-\alpha/2}^2)}$$

$$\text{Upper Limit} = \frac{(2n\hat{p} + z_{1-\alpha/2}^2 + 1) + z_{1-\alpha/2} \sqrt{z_{1-\alpha/2}^2 + \{2 - (1/n)\} + 4\hat{p}\{n(1-\hat{p}) - 1\}}}{2(n + z_{1-\alpha/2}^2)}$$

One-sided limits may be obtained by replacing  $\alpha/2$  by  $\alpha$ .

### Simple Asymptotic

The simple asymptotic formula is based on the normal approximation to the binomial distribution. The approximation is close only for very large sample sizes. The  $100(1-\alpha)\%$  confidence limits are given by

$$\hat{p} \pm z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

One-sided limits may be obtained by replacing  $\alpha/2$  by  $\alpha$ .

### Simple Asymptotic with Continuity Correction

This formula is identical to the previous one, but with continuity correction. The  $100(1-\alpha)\%$  confidence limits are

$$\left( \hat{p} - z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} - \frac{1}{2n}, \hat{p} + z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} + \frac{1}{2n} \right)$$

One-sided limits may be obtained by replacing  $\alpha/2$  by  $\alpha$ .

## Interval Widths (One-Sided vs. Two-Sided)

For two-sided intervals, the distance from the sample proportion to each of the limits may be different. Thus, instead of specifying the distance to the limits we specify the width of the interval,  $W$ .

The basic equation for determining sample size for a two-sided interval when  $W$  has been specified is

$$W = U - L$$

For one-sided intervals, the distance from the sample proportion to limit,  $D$ , is specified.

The basic equation for determining sample size for a one-sided upper limit when  $D$  has been specified is

$$D = U - \hat{p}$$

## Confidence Intervals for One Proportion

The basic equation for determining sample size for a one-sided lower limit when  $D$  has been specified is

$$D = \hat{p} - L$$

Each of these equations can be solved for any of the unknown quantities in terms of the others.

---

## Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

---

### Design Tab

The Design tab contains most of the parameters and options that you will be concerned with.

---

#### Solve For

##### Solve For

This option specifies the parameter to be solved for from the other parameters.

---

#### Confidence Interval Method

##### Confidence Interval Formula

Specify the formula to be in used in calculation of confidence intervals.

- **Exact (Clopper-Pearson)**  
The exact formula uses the binomial probabilities directly.
- **Score (Wilson)**  
This formula is based on inverting a critical ratio test.
- **Score (Continuity Correction)**  
This formula is based on inverting a critical ratio test with continuity correction.
- **Simple Asymptotic**  
The simple asymptotic formula is based on the normal approximation to the binomial.
- **Simple Asymptotic (Continuity Correction)**  
This formula is based on the normal approximation to the binomial with continuity correction.

---

#### One-Sided or Two-Sided Interval

##### Interval Type

Specify whether the interval to be used will be a two-sided confidence interval, an interval that has only an upper limit, or an interval that has only a lower limit.

## Confidence Intervals for One Proportion

---

### Confidence

#### Confidence Level

The confidence level,  $1 - \alpha$ , has the following interpretation. If thousands of samples of  $n$  items are drawn from a population using simple random sampling and a confidence interval is calculated for each sample, the proportion of those intervals that will include the true population proportion is  $1 - \alpha$ .

Often, the values 0.95 or 0.99 are used. You can enter single values or a range of values such as *0.90, 0.95 or 0.90 to 0.99 by 0.01*.

---

### Sample Size

#### N (Sample Size)

Enter one or more values for the sample size. This is the number of individuals selected at random from the population to be in the study.

You can enter a single value or a range of values.

---

### Precision

#### Confidence Interval Width (Two-Sided)

This is the distance from the lower confidence limit to the upper confidence limit.

You can enter a single value or a list of values. The value(s) must be between 0 and 1.

#### Distance from P to Limit (One-Sided)

This is the distance from the sample proportion to the lower or upper limit of the confidence interval, depending on whether the Interval Type is set to Lower Limit or Upper Limit.

You can enter a single value or a list of values. The value(s) must be between 0 and 1.

---

### Proportion

#### P (Proportion)

Enter an estimate of the proportion. The sample size and width calculations assume that the value entered here is the proportion estimate that is obtained from the sample. If the sample proportion is different from the one specified here, the width may be narrower or wider than specified.

You can enter a range of values such as *.1 .2 .3 or .1 to .5 by .1*.

## Confidence Intervals for One Proportion

## Example 1 – Calculating Sample Size

Suppose a study is planned in which the researcher wishes to construct a two-sided 95% exact (Clopper-Pearson) confidence interval for the population proportion such that the width of the interval is no wider than 0.06. The anticipated proportion estimate is 0.3, but a range of values from 0.1 to 0.5 will be included to determine the effect of the proportion estimate on necessary sample size. Instead of examining only the interval width of 0.06, widths of 0.04 and 0.10 will also be considered.

The goal is to determine the necessary sample size.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Confidence Intervals for One Proportion** procedure window by expanding **Proportions**, then **One Proportion**, then clicking on **Confidence Interval**, and then clicking on **Confidence Intervals for One Proportion**. You may make the appropriate entries as listed below or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Confidence Interval Formula .....	<b>Exact (Clopper-Pearson)</b>
Interval Type .....	<b>Two-Sided</b>
Confidence Level .....	<b>0.95</b>
Confidence Interval Width (Two-Sided) ..	<b>0.04 0.06 0.10</b>
P (Proportion) .....	<b>0.1 to 0.5 by 0.1</b>

### Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

### Numeric Results

Numeric Results for Two-Sided Confidence Intervals for One Proportion  
Confidence Interval Formula: Exact (Clopper-Pearson)

Confidence Level	Sample Size (N)	Target Width	Actual Width	Proportion (P)	Lower Limit	Upper Limit	Width if P = 0.5
0.950	914	0.040	0.040	0.100	0.081	0.121	0.066
0.950	1585	0.040	0.040	0.200	0.181	0.221	0.050
0.950	2065	0.040	0.040	0.300	0.280	0.320	0.044
0.950	2353	0.040	0.040	0.400	0.380	0.420	0.041
0.950	2449	0.040	0.040	0.500	0.480	0.520	0.040
0.950	417	0.060	0.060	0.100	0.073	0.133	0.098
0.950	715	0.060	0.060	0.200	0.171	0.231	0.075
0.950	928	0.060	0.060	0.300	0.271	0.331	0.065
0.950	1056	0.060	0.060	0.400	0.370	0.430	0.061
0.950	1098	0.060	0.060	0.500	0.470	0.530	0.060
0.950	158	0.100	0.100	0.100	0.058	0.158	0.161
0.950	264	0.100	0.100	0.200	0.153	0.253	0.124
0.950	341	0.100	0.100	0.300	0.252	0.352	0.109
0.950	387	0.100	0.100	0.400	0.351	0.451	0.102
0.950	402	0.100	0.100	0.500	0.450	0.550	0.100

## Confidence Intervals for One Proportion

### References

- Fleiss, J. L., Levin, B., Paik, M.C. 2003. *Statistical Methods for Rates and Proportions*. Third Edition. John Wiley & Sons. New York.
- Newcombe, R. G. 1998. 'Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods.' *Statistics in Medicine*, 17, pp. 857-872.

### Report Definitions

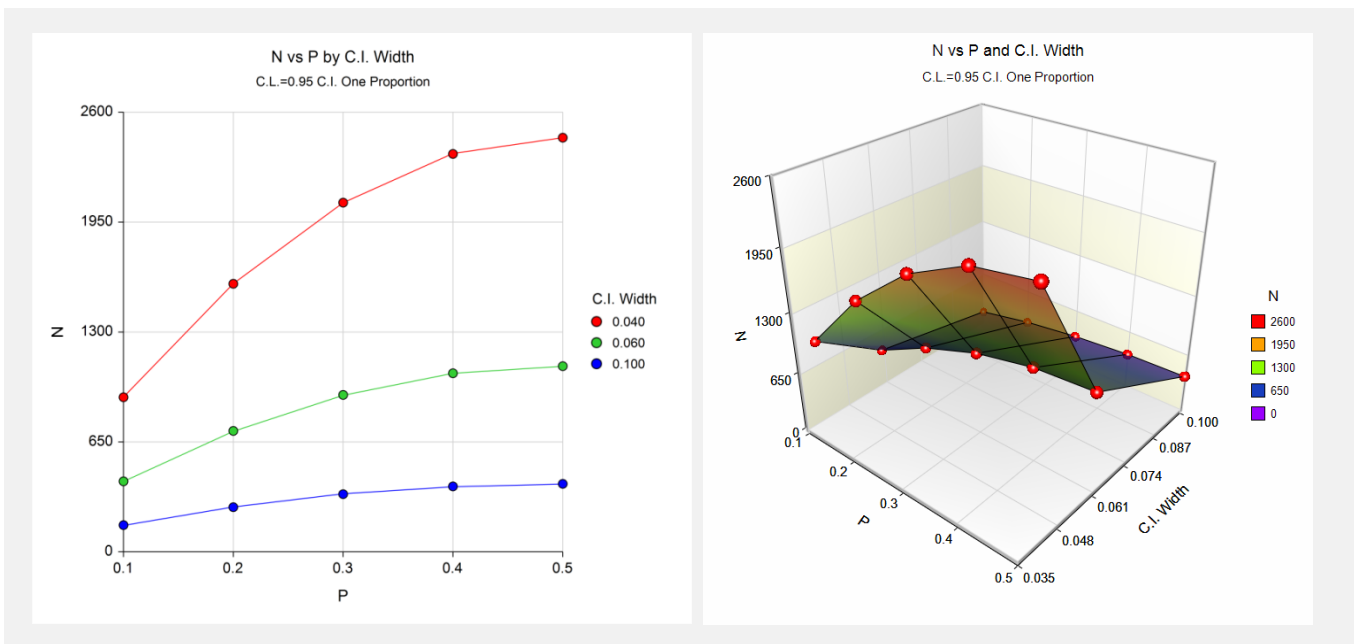
- Confidence level is the proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that would contain the population proportion.
- N is the size of the sample drawn from the population.
- Width is the distance from the lower limit to the upper limit.
- Target Width is the value of the width that is entered into the procedure.
- Actual Width is the value of the width that is obtained from the procedure.
- Proportion (P) is the assumed sample proportion.
- Lower Limit is the lower limit of the confidence interval.
- Upper Limit is the upper limit of the confidence interval.
- Width if P = 0.5 is the maximum width for a confidence interval with sample size N.

### Summary Statements

A sample size of 914 produces a two-sided 95% confidence interval with a width equal to 0.040 when the sample proportion is 0.100.

This report shows the calculated sample size for each of the scenarios.

## Plots Section



These plots show the sample size versus the sample proportion for the three confidence interval widths.

## Confidence Intervals for One Proportion

**Example 2 – Validation using Fleiss, Levin, and Paik**

Fleiss, Levin, and Paik (2003), pages 22-23, give an example of a calculation for an exact (Clopper-Pearson) one-sided lower limit confidence interval for a single proportion when the confidence level is 95%, the sample proportion is 0.92, and the distance from the lower limit to the sample proportion is 0.15104. The necessary sample size is 25.

**Setup**

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Confidence Intervals for One Proportion** procedure window by expanding **Proportions**, then **One Proportion**, then clicking on **Confidence Interval**, and then clicking on **Confidence Intervals for One Proportion**. You may make the appropriate entries as listed below or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Confidence Interval Formula .....	<b>Exact (Clopper-Pearson)</b>
Interval Type .....	<b>Lower Limit</b>
Confidence Level .....	<b>0.95</b>
Distance from P to Limit (One-Sided) .....	<b>0.15104</b>
P (Proportion) .....	<b>0.92</b>

**Output**

Click the Calculate button to perform the calculations and generate the following output.

**Numeric Results**

Confidence Level	Sample Size (N)	Target Distance from P to Lower Limit	Actual Distance from P to Lower Limit	Proportion (P)	Lower Limit	Upper Limit	Distance from P to Limit if P = 0.5
0.950	25	0.151	0.151	0.920	0.769	1	0.177

PASS also calculated the necessary sample size to be 25.

## Confidence Intervals for One Proportion

**Example 3 – Validation using Newcombe**

Newcombe (1998a), pages 860-861, gives an example of a calculation for a two-sided confidence interval for a single proportion for each of the methods when the confidence level is 95%. Here we validate the score method with continuity correction. The sample proportion is 0.034483, and the interval width is 0.1945. The necessary sample size is 29.

**Setup**

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Confidence Intervals for One Proportion** procedure window by expanding **Proportions**, then **One Proportion**, then clicking on **Confidence Interval**, and then clicking on **Confidence Intervals for One Proportion**. You may make the appropriate entries as listed below or open **Example 3** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Confidence Interval Formula .....	<b>Score (Contin. Correction)</b>
Interval Type .....	<b>Two-Sided</b>
Confidence Level .....	<b>0.95</b>
Confidence Interval Width (Two-Sided) ..	<b>0.1945</b>
P (Proportion) .....	<b>0.034483</b>

**Output**

Click the Calculate button to perform the calculations and generate the following output.

**Numeric Results**

Confidence Level	Sample Size (N)	Target Width	Actual Width	Proportion (P)	Lower Limit	Upper Limit	Width if P = 0.5
0.950	29	0.195	0.194	0.034	0.002	0.196	0.372

PASS also calculated the necessary sample size to be 29.



## Confidence Intervals for One Proportion

**Example 4 – Zero Events, Validation using Lachin**

Lachin (2000), page 19, gives an example of a calculation for a one-sided upper limit exact confidence interval for a single proportion when the confidence level is 95%, the sample proportion is 0, and the upper bound is 0.01. The necessary sample size is 299.

**Setup**

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Confidence Intervals for One Proportion** procedure window by expanding **Proportions**, then **One Proportion**, then clicking on **Confidence Interval**, and then clicking on **Confidence Intervals for One Proportion**. You may make the appropriate entries as listed below or open **Example 4** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Confidence Interval Formula .....	<b>Exact (Clopper-Pearson)</b>
Interval Type .....	<b>Upper Limit</b>
Confidence Level.....	<b>0.95</b>
Distance from P to Limit (One-Sided).....	<b>0.01</b>
P (Proportion) .....	<b>0</b>

**Output**

Click the Calculate button to perform the calculations and generate the following output.

**Numeric Results**

Confidence Level	Sample Size (N)	Target Distance from P to Upper Limit	Actual Distance from P to Upper Limit	Proportion (P)	Lower Limit	Upper Limit	Distance from P to Limit if P = 0.5
0.950	299	0.010	0.010	0.000	0	0.010	0.049

*PASS* also calculated the necessary sample size to be 299.