

Chapter 116

Confidence Intervals for One Proportion from a Finite Population

Introduction

This routine calculates the sample size necessary to achieve a specified interval width at a stated confidence level for a confidence interval of one proportion estimated from a finite population.

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

Confidence Interval Formulas

We use the results of Machin, Campbell, Tan, and Tan (2009). Let \hat{p} be the sample proportion, r the number of successes in a sample of size n , N the population size, α the value of $1 - \text{confidence level}$, and $\hat{p} = r/n$. The asymptotic formula for a $100(1-\alpha)\%$ confidence interval of p based on the normal approximation to the hypergeometric distribution is:

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

Confidence Intervals for One Proportion from a Finite Population

Hence, the equation relating d (the precision, margin of error, or half-width) is

$$d = z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})(N-n)}{nN}}$$

This equation can be rearranged to obtain the precision, confidence level ($1 - \alpha$), or sample size (n).

The confidence level, $1 - \alpha$, has the following interpretation. If thousands of samples of n items are drawn from a population of N items 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 $1 - \alpha$.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Data 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.

Population Size

N (Population Size)

Enter the total number of subjects in the population. Of course, N must be greater than n .

Sample Size

n (Sample Size)

Enter the value of for the sample size, n . This is the number of individuals selected from the population.

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

Proportion

P (Sample Proportion)

Enter the anticipated sample proportion. You can enter a single value or a range of values such as

0.1, 0.2, 0.3

or

0.1 to 0.5 by 0.1.

You must be careful that the values of P and d are compatible. That is, $0 < P - d < P + d < 1$.

Confidence Intervals for One Proportion from a Finite Population

Precision

d (Precision, Half Width)

Enter d , the precision, margin of error, or interval half width.

You can enter a single value or a list of values. You must be careful that the values of P and d are compatible. That is, $0 < P - d < P + d < 1$.

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 a single value or a range of values such as

0.90, 0.95

or

0.90 to 0.99 by 0.01.

Confidence Intervals for One Proportion from a Finite Population

Example 1 – Finding Sample Size

Suppose a study is planned in which the researcher wishes to construct a two-sided 95% confidence interval for the population proportion in a population of 3000 items. The width of the interval is to be no wider than 0.06. The anticipated proportion estimate is 0.3, but a range of values from 0.2 to 0.4 will be investigated 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 0.08 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 from a Finite Population** procedure window by expanding **Proportions**, then **One Proportion**, then clicking on **Confidence Interval**, and then clicking on **Confidence Intervals for One Proportion from a Finite Population**. You may then 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>
Design Tab	
Solve For	Sample Size
N (Population Size)	3000
P (Sample Proportion)	0.2 to 0.4 by 0.05
d (Precision, Half Width)	0.04 to 0.08 by 0.02
Confidence Level	0.95

Annotated Output

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

Numeric Results

Confidence Level	Sample Size (n)	Precision or Half-Width (d)	Sample Proportion (P)	Population Size (N)
0.950	341	0.0400	0.2000	3000
0.950	392	0.0400	0.2500	3000
0.950	432	0.0400	0.3000	3000
0.950	463	0.0400	0.3500	3000
0.950	484	0.0400	0.4000	3000
0.950	162	0.0600	0.2000	3000
0.950	188	0.0600	0.2500	3000
0.950	209	0.0600	0.3000	3000
0.950	225	0.0600	0.3500	3000
0.950	236	0.0600	0.4000	3000
0.950	94	0.0800	0.2000	3000
0.950	109	0.0800	0.2500	3000
0.950	121	0.0800	0.3000	3000
0.950	131	0.0800	0.3500	3000
0.950	138	0.0800	0.4000	3000

Confidence Intervals for One Proportion from a Finite Population

References

Machin, D., Campbell, M., Tan, S.B., and Tan, S.H. 2009. Sample Size Tables for Clinical Studies, 3rd Edition. Wiley-Blackwell. Chichester, UK.

Report Definitions

d is the precision, margin of error, or half-width of the two-sided confidence interval.

Confidence Level is the proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that contain the population proportion.

n is the size of the sample drawn from the N members of population.

P is the anticipated value of proportion.

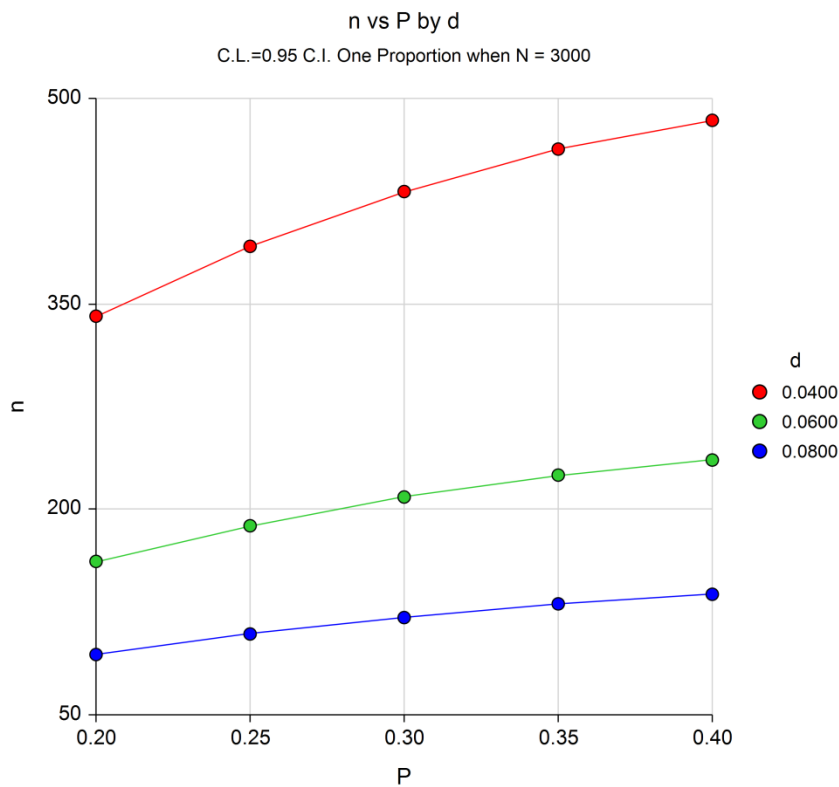
N is the size of the population.

Summary Statements

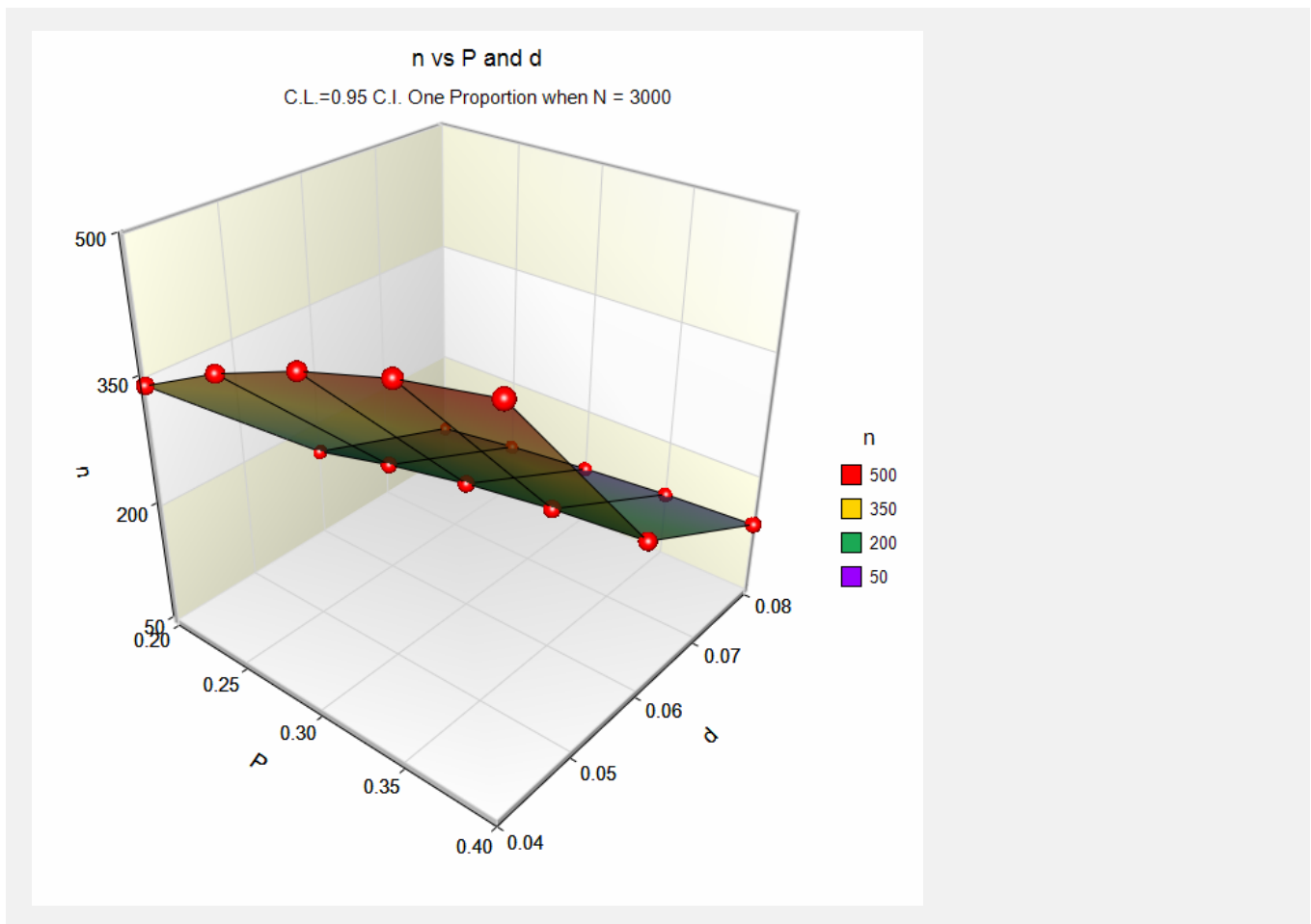
A sample size of 341 from a population of 3000 produces a two-sided 95% confidence interval with a precision (half-width) of 0.0400 when the actual proportion is near 0.2000.

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

Plots Section



Confidence Intervals for One Proportion from a Finite Population



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

Confidence Intervals for One Proportion from a Finite Population

Example 2 – Validation using Machin et al. (2009)

Machin et al. (2009), page 132, give an example of a sample size calculation for a finite population of 1000 of a confidence interval for a single proportion when the confidence level is 95%, the sample proportion is 0.2 and the precision is 0.04. They determine the required sample size to be 277.

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 from a Finite Population** procedure window by expanding **Proportions**, then **One Proportion**, then clicking on **Confidence Interval**, and then clicking on **Confidence Intervals for One Proportion from a Finite Population**. You may then 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>
Design Tab	
Solve For	Sample Size
N (Population Size).....	1000
P (Sample Proportion)	0.20
d (Precision, Half Width).....	0.04
Confidence Level.....	0.95

Output

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

Numeric Results

Confidence Level	Sample Size (n)	Precision or Half-Width (d)	Sample Proportion (P)	Population Size (N)
0.950	278	0.0400	0.2000	1000

PASS also calculated the necessary sample size to be 278 which is within rounding of the 277 they found. If you check, the sample size of 277 results in a precision of 0.0401 which is slightly larger than the required 0.0400. This is why PASS reported 278.