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 – confidence level, and $\hat{p} = r\ /\ n$. The asymptotic formula for a 100(1 - α)% confidence interval of *p* based on the normal approximation to the hypergeometric distribution is:

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

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 – α), or sample size (*n*).

The confidence level, 1 – α, 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 – α.

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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

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**

Output

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

Numeric Reports

Numeric Results

Solve For: Sample Size

Confidence Level	Sample Size n	Precision or Half-Width d	Sample Proportion P	Population Size N
0.95	341	0.04	0.20	3000
0.95	392	0.04	0.25	3000
0.95	432	0.04	0.30	3000
0.95	463	0.04	0.35	3000
0.95	484	0.04	0.40	3000
0.95	162	0.06	0.20	3000
0.95	188	0.06	0.25	3000
0.95	209	0.06	0.30	3000
0.95	225	0.06	0.35	3000
0.95	236	0.06	0.40	3000
0.95	94	0.08	0.20	3000
0.95	109	0.08	0.25	3000
0.95	121	0.08	0.30	3000
0.95	131	0.08	0.35	3000
0.95	138	0.08	0.40	3000

Confidence Level The proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that contain the population proportion.
n The size of the sample drawn from the N members of population.
d The precision, margin of error, or half-width of the two-sided confidence interval.
P The anticipated value of the sample proportion.
N The size of the population.

Summary Statements

A single-group design will be used to obtain a two-sided 95% confidence interval for a single proportion. The sample proportion is assumed to be 0.2. To produce a confidence interval with a half-width of no more than 0.04, 341 subjects from a population of 3000 subjects will be needed.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size n	Dropout-Inflated Enrollment Sample Size n'	Expected Number of Dropouts D
20%	341	427	86
20%	392	490	98
20%	432	540	108
20%	463	579	116
20%	484	605	121
20%	162	203	41
20%	188	235	47
20%	209	262	53
20%	225	282	57
20%	236	295	59
20%	94	118	24
20%	109	137	28
20%	121	152	31
20%	131	164	33
20%	138	173	35

Dropout Rate The percentage of subjects (or items) that are expected to be lost at random during the course of the study and for whom no response data will be collected (i.e., will be treated as "missing"). Abbreviated as DR.
n The evaluable sample size at which the confidence interval is computed. If n subjects are evaluated out of the n' subjects that are enrolled in the study, the design will achieve the stated confidence interval.
n' The total number of subjects that should be enrolled in the study in order to obtain n evaluable subjects, based on the assumed dropout rate. After solving for n, n' is calculated by inflating n using the formula n' = n / (1 - DR), with n' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. (2018) pages 32-33.)
D The expected number of dropouts. D = n' - n.

Dropout Summary Statements

Anticipating a 20% dropout rate, 427 subjects should be enrolled to obtain a final sample size of 341 subjects.

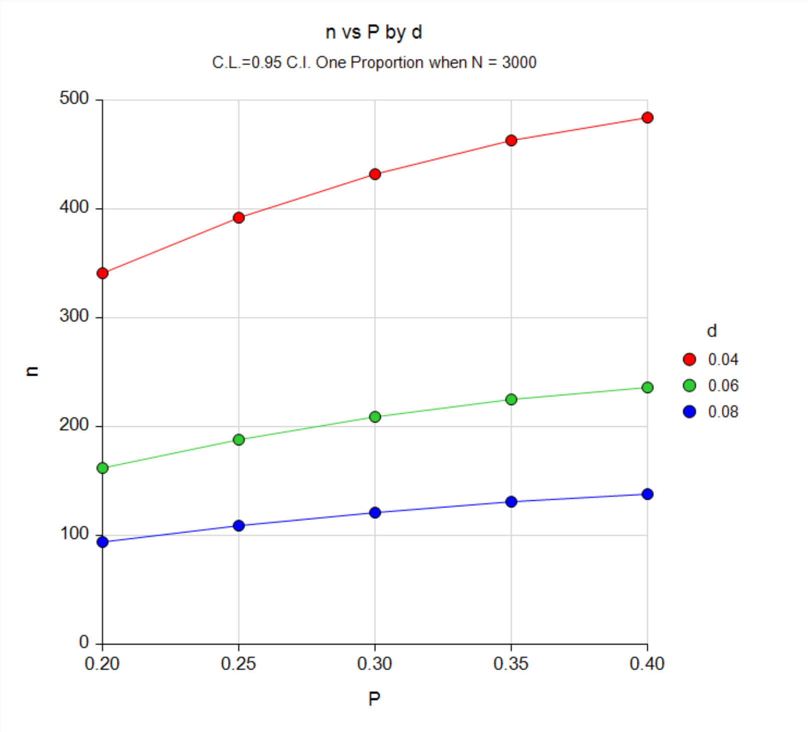
References

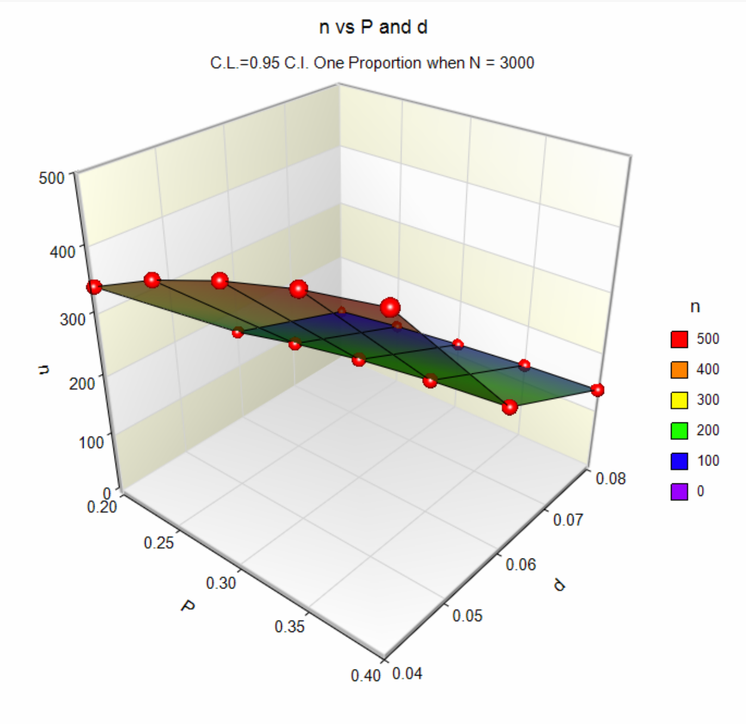
Machin, D., Campbell, M., Tan, S.B., and Tan, S.H. 2009. Sample Size Tables for Clinical Studies, 3rd Edition. Wiley-Blackwell. Chichester, UK.

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

Plots Section

Plots





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

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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

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

Solve For: Sample Size

Confidence Level	Sample Size n	Precision or Half-Width d	Sample Proportion P	Population Size N
0.95	278	0.04	0.2	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.