Chapter 653

# Confidence Intervals for One Variance using Relative Error

## Introduction

This routine calculates the necessary sample size such that a sample variance estimate will achieve a specified relative distance from the true population variance at a stated confidence level when the underlying data distribution is normal.

Caution: This procedure controls the relative width of the interval as a proportion of the true population variance. For controlling the absolute width of the interval see the procedures Confidence Intervals for One Variance using Variance and Confidence Intervals for One Variance with Tolerance Probability.

## Technical Details

Following the results of Desu and Raghavarao (1990) and Greenwood and Sandomire (1950), let $s^2$ be the variance estimate based on a sample from a normal distribution with unknown $\mu$ and unknown $\sigma^2$. Let $r$ be the proportion of $\sigma^2$ such that $s^2$ is within $r\sigma^2$ of $\sigma^2$ with desired confidence. If

$$p_1 = \Pr(s^2 > \sigma^2 + r\sigma^2) = \Pr(s^2 > \sigma^2(1 + r))$$

and

$$p_2 = \Pr(s^2 < \sigma^2 - r\sigma^2) = \Pr(s^2 < \sigma^2(1 - r))$$

The confidence level for estimating $\sigma^2$ within proportion $r$ is $1 - p_1 - p_2$.

Since $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$, it is useful to rewrite $p_1$ and $p_2$ as

$$p_1 = \Pr\left(\frac{(n-1)s^2}{\sigma^2} > (n-1)(1+r)\right)$$

and

$$p_2 = \Pr\left(\frac{(n-1)s^2}{\sigma^2} < (n-1)(1-r)\right)$$

Using the chi-square distribution, these equations can be solved for any of the unknown quantities ($n$, $r$, $p_1 + p_2$) in terms of the others.

## 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 the variance estimates are obtained from each sample, the proportion of those estimates that are within $r\sigma^2$ of $\sigma^2$ is $1 - \alpha$.

# Example 1 – Calculating Sample Size

Suppose a study is planned in which the researcher wishes to be 95% confident that estimated variance is within 10% of the true population variance. In addition to 10% relative error, 5%, 15%, 20% and 25% 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**
Interval Type ................................................. **Two-Sided**
Confidence Level (1 - Alpha) ........................ **0.95**
Relative Error ................................................ **0.05 to 0.25 by 0.05**

## Output

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

### Numeric Reports

**Numeric Results**

Solve For: Sample Size
Interval Type: Two-Sided

| Confidence Level Target | Confidence Level Actual | Sample Size N | Relative Error |
|---|---|---|---|
| 0.95 | 0.95 | 3074 | 0.05 |
| 0.95 | 0.95 | 769 | 0.10 |
| 0.95 | 0.95 | 342 | 0.15 |
| 0.95 | 0.95 | 192 | 0.20 |
| 0.95 | 0.95 | 123 | 0.25 |

Confidence Level The proportion of variance estimates that will be within the relative error of the true variance.
Target Confidence Level The value of the confidence level that is entered into the procedure.
Actual Confidence Level The value of the confidence level that is obtained from the procedure.
N The size of the sample drawn from the population.
Relative Error The distance from the true variance as a proportion of the true variance.

### Summary Statements

A single-group design will be used to obtain an estimate of the variance with relative precision. The Chi-square-based methods of Desu and Raghavarao (1990) and Greenwood and Sandomire (1950) will be used. The underlying data distribution is assumed to be Normal. To obtain an estimate of the variance where the probability is 0.95 (95% confidence) that the estimate of the variance will be within 5% of the true population variance (two-sided), 3074 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% | 3074 | 3843 | 769 |
| 20% | 769 | 962 | 193 |
| 20% | 342 | 428 | 86 |
| 20% | 192 | 240 | 48 |
| 20% | 123 | 154 | 31 |

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, 3843 subjects should be enrolled to obtain a final sample size of 3074 subjects.

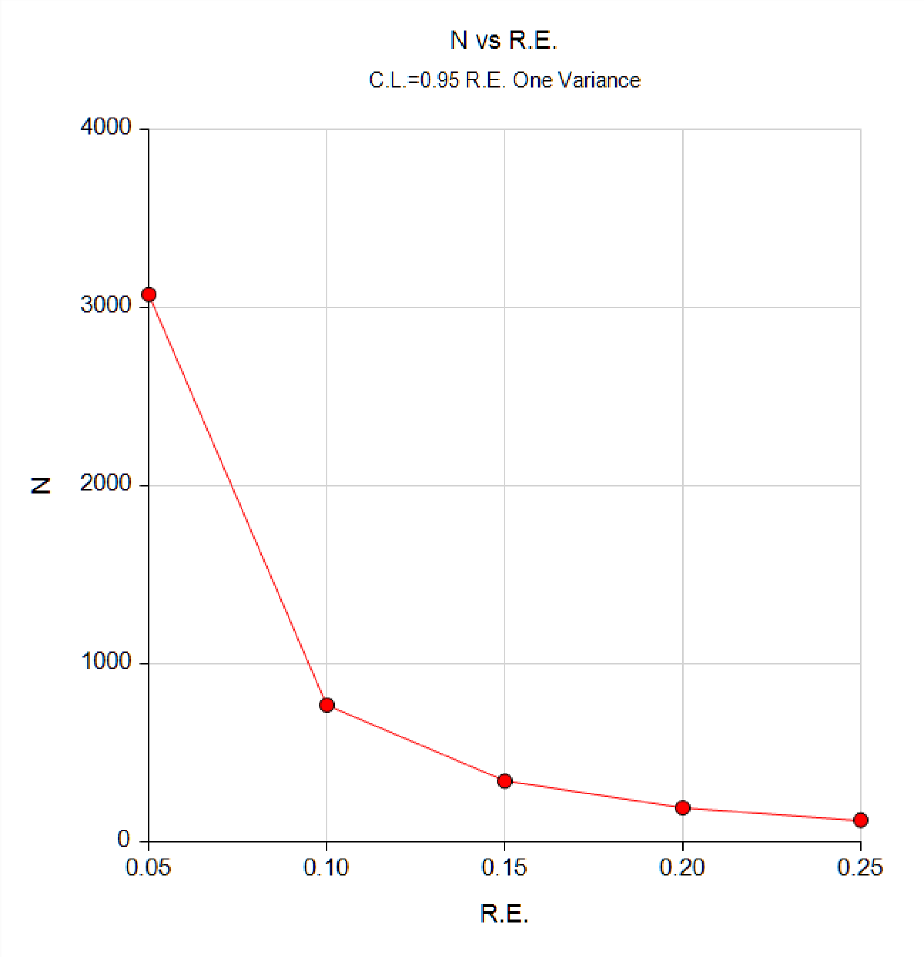
### References

Desu, M. M. and Raghavarao, D. 1990. Sample Size Methodology. Academic Press. New York.
Greenwood, J. A. and Sandomire, M. M. 1950. 'Sample Size Required for Estimating the Standard Deviation as a Per Cent of its True Value', Journal of the American Statistical Association, Vol. 45, No. 250, pp. 257-260.

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

## Plots Section

**Plots**




This plot shows the sample size versus the relative error.

# Example 2 – Validation using Desu and Raghavarao (1990)

Desu and Raghavarao (1990) page 6 give an example of a calculation in which the desired confidence level is 95% and the relative error is 20%. This calculation is based on a large sample approximation formula. The necessary sample size is 194.

## 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**
Interval Type ................................................ **Two-Sided**
Confidence Level (1 - Alpha) ........................ **0.95**
Relative Error ................................................ **0.2**

## Output

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

**Numeric Results**

Solve For: Sample Size
Interval Type: Two-Sided

| Confidence Level | | Sample | |
|---|---|---|---|
| | | Size | Relative |
| Target | Actual | N | Error |
| 0.95 | 0.95 | 192 | 0.2 |

**PASS** calculated the necessary sample size to be 192 but did not use the approximation formulas.

Using direct calculation with the non-approximate chi-square formulas with a sample size of 194, the confidence level is 1 – (0.0300164564 + 0.0188126383) = 0.951171.

Using direct calculation with the non-approximate chi-square formulas with a sample size of 192, the confidence level is 1 – (0.0306634808 + 0.0193315700) = 0.950005, which is closer to the prescribed confidence level. Thus, 192 is the correct value.