Chapter 817

Confidence Intervals for Intraclass Correlation

Introduction

This routine calculates the sample size needed to obtain a specified width of an intraclass correlation coefficient confidence interval at a stated confidence level.

The **intraclass correlation** coefficient (ρ in this chapter) is the product-moment correlation calculated among observations on the same subject. The intraclass correlation analyzed in this procedure comes from a one-way random effects ANOVA model, a two-way random-effects ANOVA model, and a two-way mixed effects ANOVA model.

Technical Details

Shrout and Fleiss (1979) and Bonett (2002) present the formulas for constructing $100(1 - \alpha)\%$ confidence intervals for the intraclass correlation coefficient (ICC). We will adopt their notation as we present these formulas.

Suppose that each of N subjects (targets) yields K observations. These observations may be from different raters, instruments, or other measurement mechanisms. Such data may be analyzed using one of three different ANOVA models.

Model 1: One-Way Random Effects Model

The raters of each subject are selected as a random sample from the (larger) population of raters. A different set of K raters is selected for each subject. Raters are nested in subjects. The ANOVA model is

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \text{ where } \alpha_i \sim N(0, \sigma_\alpha^2) \text{ and } \varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2).$$

Model 2: Two-Way Random Effects Model

A set of K raters is randomly selected from the population of raters. These raters are then used to judge all subjects. The ANOVA model is

$$Y_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ij} \text{ where } \alpha_i \sim N(0, \sigma_\alpha^2), \beta_j \sim N(0, \sigma_\beta^2), \alpha\beta_{ij} \sim N(0, \sigma_{\alpha\beta}^2), \text{ and } \varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2).$$

Model 3: Two-Way Fixed Effects Model

The K raters are the only raters of interest. No reference is given to a population of raters. These raters are then used to judge all subjects. The ANOVA model is

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} \text{ where } \alpha_i \sim N(0, \sigma_\alpha^2) \text{ and } \varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2).$$

Estimation of ρ

The intraclass correlation coefficient is defined as

$$\rho = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2}$$

The ICC ρ is estimated from the mean squares of the ANOVA table as follows

$$r = \frac{MS_B - MS_E}{MS_B + (K-1)MS_E}$$

where MS_B is the between-subject mean square, MS_E is the within-subject mean square in the one-way model and the residual mean square in both the two-way models.

Confidence limits r_L and r_U for ρ are obtained using the formulas

$$r_L = \frac{F_L - 1}{F_L + K - 1}, \qquad r_U = \frac{F_U - 1}{F_U + K - 1}$$

where

$$F_L = \frac{F_O}{F_{1-\alpha/2,V2,V1}}, \quad F_U = F_O F_{1-\alpha/2,V1,V2}, \quad F_O = \frac{MS_B}{MS_E}, \quad V1 = (N-b)(K-1), \quad V2 = N-1.$$

Here, b is zero for the one-way design and one for the two-way designs.

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

For a particular value of r, these confidence limits may be calculated and the interval width $r_U - r_L$ may be compared to a desired value.

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 correlation is $1 - \alpha$.

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.

Model and Confidence Interval

ANOVA Model Type

Specify the type of ANOVA model to be used. The possible choices are One-Way Random Effects, Two-Way Random Effects, or Two-Way Mixed Effects.

Interval Type

Specify whether the confidence interval for the population correlation is two-sided or one-sided. A one-sided interval is often called a **confidence bound** rather than a confidence interval because it only has one limit.

Two-Sided

The two-sided confidence interval is defined by two limits: an upper confidence limit (UCL) and a lower confidence limit (LCL).

These limits are constructed so that the designated proportion (confidence level) of such intervals will include the true population value.

Upper One-Sided

The upper confidence interval (or bound) is defined by a limit above the estimated parameter value. The limit is constructed so that the designated proportion (confidence level) of such limits has the true population value below them.

Lower One-Sided

The lower confidence interval (or bound) is defined by a limit below the estimated parameter value. The limit is constructed so that the designated proportion (confidence level) of such limits has the true population value above them.

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 correlation 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 (Number of Subjects)

Enter one or more values for N, the number of subjects. This is the number of experimental units selected at random from the population.

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

K (Observations per Subject)

Enter one or more values for K, the number of observations per subject. This is the number of scores made by judges or raters.

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

Precision

Width of Confidence Interval

This is the distance between the lower confidence limit (LCL) and the upper confidence limit (UCL). Its calculation is UCL - LCL. It is a measure of the precision of the confidence interval.

Since the ICC is usually positive, this width usually ranges between 0 and 1.

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

Distance from r to Limit

This is the distance from r to the lower confidence limit (LCL) or the upper confidence limit (UCL). It is calculated using |r-LCL| or |UCL-r|. The range is between 0 and 1.

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

Estimated Sample Intraclass Correlation

r (Sample Intraclass Correlation)

Enter a planning estimate of the intraclass correlation. The sample size and width calculations assume that the value entered here is the ICC that is obtained from the sample. This is sometimes referred to as a *planning estimate*.

The range of the values of the sample correlation that can be entered is 0 to 1.

You can enter a range of values such as 0.1 0.3 0.5 or 0.1 to 0.5 by .2.

Example 1 – Calculating Sample Size

Suppose a study is planned that will be analyzed with a two-way mixed-effects ANOVA model in which the researcher wishes to construct a two-sided 95% confidence interval for the intraclass correlation such that the width of the interval is no wider than 0.1. The researcher would like to examine values of K from 2 to 5 with ICC = 0.50. The goal is to determine the necessary sample size, N.

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 Intraclass Correlation** procedure window by expanding **Correlation**, then **Correlation**, then clicking on **Confidence Interval**, and then clicking on **Confidence Intervals for Intraclass Correlation**. 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**.

Option   **Value**

Design Tab
Solve For .... **Sample Size**
ANOVA MODEL Type .... **Two-Way Mixed Effects**
Interval Type .... **Two-Sided**
Confidence Level (1 – Alpha) **0.95**
K (Observations per Subject) .... **2 to 5 by 1**
Width of Confidence Interval **0.1**
r (Sample Intraclass Correlation)............ **0.5**

Annotated Output

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

Numeric Results

Numeric Results for Two-Sided Confidence Interval for Intraclass Correlation in a Two-Way Design

Confidence Level	Number of Subjects N	Observations per Subject K	Width UCL-LCL	Sample Intraclass Correlation r	Lower Confidence Limit LCL	Upper Confidence Limit UCL
0.950	866	2	0.100	0.500	0.448	0.548
0.950	513	3	0.100	0.500	0.449	0.549
0.950	401	4	0.100	0.500	0.450	0.550
0.950	346	5	0.100	0.500	0.450	0.550

Report Definitions
Confidence Level is the proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that would contain the true correlation.
Number of Subjects N is the size of the random sample of subjects drawn from the population.
Observations per Subject K is the number of measurements that are made on each subject.
Width (UCL-LCL) is the width of the confidence interval. It is the distance from the lower limit to the upper limit.
Sample Intraclass Correlation r is the estimate of intraclass correlation.
Lower and Upper Confidence Limits are the actual limits that would result from a dataset with these statistics.

References
Bonett, D. G. 2002. 'Sample size requirements for estimating intraclass correlations with desired precision.'
    Statistics in Medicine, Vol 21, 1331-1335.
Shrout, P. E. and Fleiss, J. L. 1979. 'Intraclass Correlations: Uses in Assessing Rater Reliability.'
    Psychological Bulletin, Vol 86, No. 2, 420-428.
Bartko, John J. 1966. 'The intraclass correlation coefficient as a measure of reliability.' Psychological
    Reports, Vol 19, 3-11.
Mathews, Paul. 2010. Sample Size Calculations: Practical Methods for Engineers and Scientists. Mathews Malnar
    and Bailey, Inc. www.mmbstatistical.com.

Summary Statements
A random sample of 866 subjects who are each measured 2 times produces a two-sided 95%
confidence interval with a width of 0.100 when the estimated intraclass correlation is 0.500.
The data will be analyzed using a two-way mixed-effects ANOVA model.

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

Plots Section
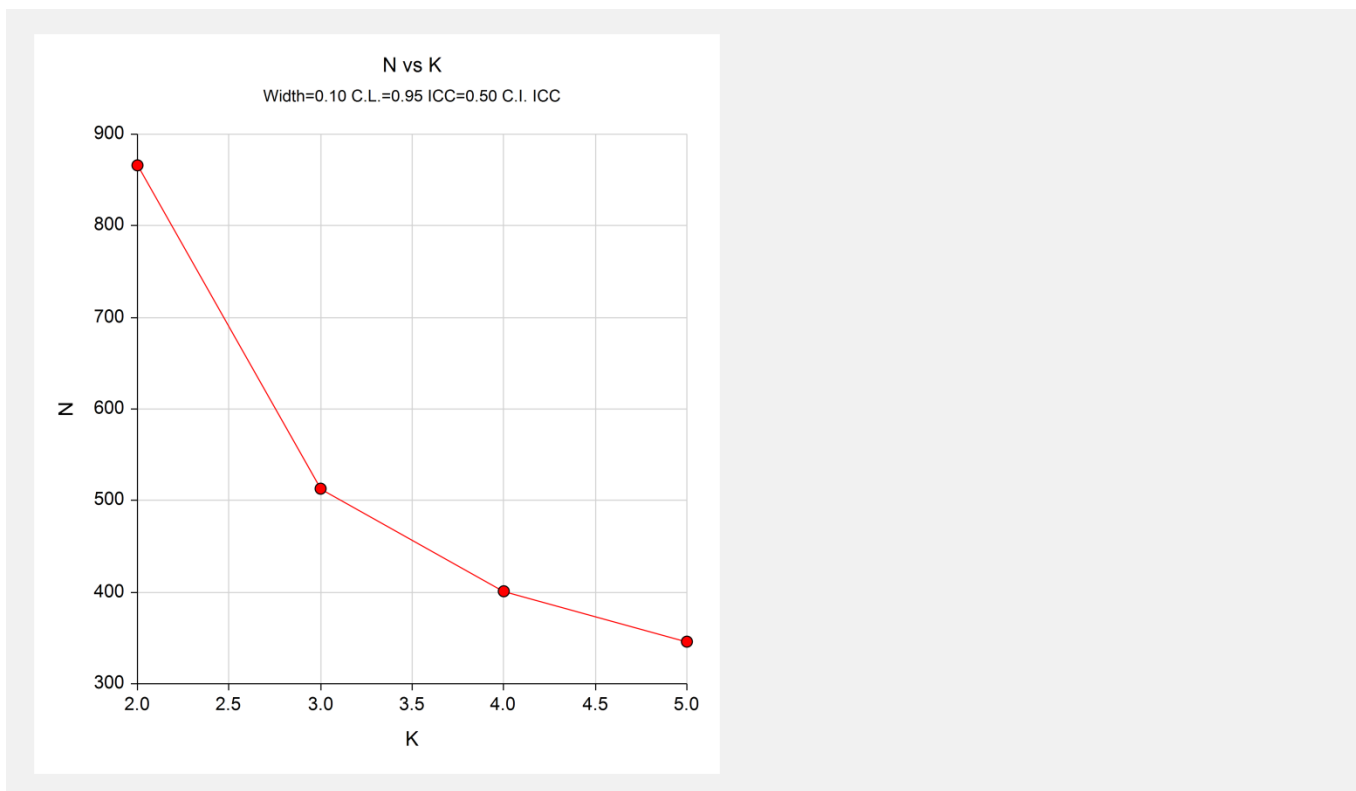
This plot shows the sample size versus the value of K.

Example 2 – Validation using Bonett

Bonett (2002), page 1334, gives example calculations of the number of subjects needed of a two-sided confidence interval of the intraclass correlation when the confidence level is 95%, the intraclass correlation is 0.30, for K equal to 2, 3, 5, and 10. His results for N were 319, 161, 92, and 59. He was using a two-way ANOVA model.

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 Intraclass Correlation** procedure window by expanding **Correlation**, then **Correlation**, then clicking on **Confidence Interval**, and then clicking on **Confidence Intervals for Intraclass Correlation**. 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**.

Option   **Value**

Design Tab
Solve For .... **Sample Size**
ANOVA MODEL Type .... **Two-Way Mixed Effects**
Interval Type .... **Two-Sided**
Confidence Level (1 – Alpha) **0.95**
K (Observations per Subject) .... **2 to 5 by 1**
Width of Confidence Interval **0.2**
r (Sample Intraclass Correlation)............ **0.3**

Annotated Output

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

Numeric Results for Two-Sided Confidence Interval for Intraclass Correlation in a Two-Way Design

Confidence Level	Number of Subjects N	Observations per Subject K	Width UCL-LCL	Sample Intraclass Correlation r	Lower Confidence Limit LCL	Upper Confidence Limit UCL
0.950	319	2	0.200	0.300	0.197	0.397
0.950	161	3	0.200	0.300	0.202	0.401
0.950	92	5	0.200	0.300	0.207	0.407
0.950	59	10	0.200	0.300	0.213	0.412

PASS matches Bonett's results exactly.