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 Mixed 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 $1 - \alpha$.

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

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**
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**

Output

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

Numeric Reports

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

Solve For: Sample Size
_____

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
0.95	866	2	0.1	0.5	0.448	0.548
0.95	513	3	0.1	0.5	0.449	0.549
0.95	401	4	0.1	0.5	0.450	0.550
0.95	346	5	0.1	0.5	0.450	0.550

Confidence Level	The proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that would contain the true correlation.
N	Number of Subjects. The size of the random sample of subjects drawn from the population.
K	Observations per Subject. The number of measurements that are made on each subject.
UCL - LCL	The width of the confidence interval. It is the distance from the lower limit to the upper limit.

r	Sample Intraclass Correlation. The estimate of intraclass correlation.
Lower and Upper Confidence Limits	The actual limits that would result from a dataset with these statistics.

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 0.1 when the estimated intraclass correlation is 0.5. The data will be analyzed using a two-way mixed-effects ANOVA model.
───

Dropout-Inflated Sample Size
───

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	866	1083	217
20%	513	642	129
20%	401	502	101
20%	346	433	87

───

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, 1083 subjects should be enrolled to obtain a final sample size of 866 subjects.
───

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.
───

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

Plots Section

Plots
_____

N vs K
Width=0.1 C.L.=0.95 ICC=0.5 C.I. ICC
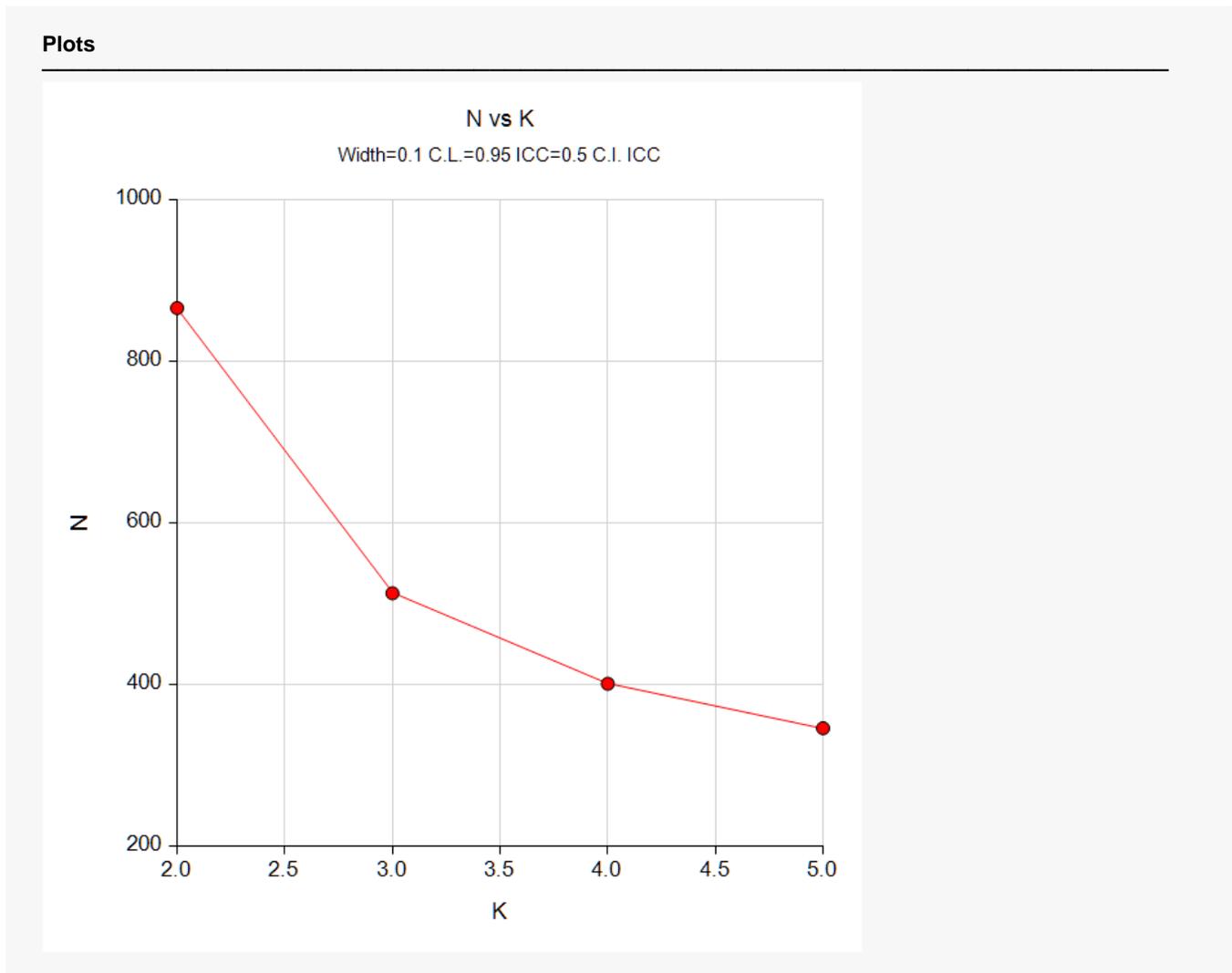
This plot shows the sample size versus the value of K.

Example 2 – Validation using Bonett (2002)

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

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**
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**

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
_____

Solve For: Sample Size
_____

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
0.95	319	2	0.2	0.3	0.197	0.397
0.95	161	3	0.2	0.3	0.202	0.401
0.95	92	5	0.2	0.3	0.207	0.407
0.95	59	10	0.2	0.3	0.213	0.412

PASS matches Bonett's results exactly.