RELIABILITY

The RELIABILITY procedure employs one of two different computing methods, depending upon the MODEL specification and options and statistics requested.

Method 1 does not involve computing a covariance matrix. It is faster than method 2 and, for large problems, requires much less workspace. However, it can compute coefficients only for ALPHA and SPLIT models, and it does not allow computation of a number of optional statistics, nor does it allow matrix input or output. Method 1 is used only when alpha or split models are requested and only FRIEDMAN, COCHRAN, DESCRIPTIVES, SCALE, and/or ANOVA are specified on the STATISTICS subcommand and/or TOTAL is specified on the SUMMARY subcommand.

Method 2 requires computing a covariance matrix of the variables. It is slower than method 1 and requires more space. However, it can process all models, statistics, and options.

The two methods differ in one other important respect. Method 1 will continue processing a scale containing variables with zero variance and leave them in the scale. Method 2 will delete variables with zero variance and continue processing if at least two variables remain in the scale. If item deletion is required, method 2 can be selected by requesting the covariance method.

Notation

There are N persons taking a test that consists of k items. A score X_{ji} is given to the jth person on the ith item.

$i \rightarrow$ items

$$
\begin{array}{c|cccccc|c}
 & 1 & 2 & \cdots & i & \cdots & k & \\
\hline
1 & X_{11} & X_{12} & & & & X_{1k} & P_1 \\
\vdots & & & & & & & \vdots \\
j & & & & X_{ji} & & & P_j \\
\vdots & & & & & & & \vdots \\
N & X_{N1} & X_{N2} & & & & X_{Nk} & P_N \\
\hline
 & T_1 & T_2 & \cdots & T_i & \cdots & T_k & G
\end{array}
$$

If the model is SPLIT, k_1 items are in part 1 and $k_2 = k - k_1$ are in part 2. If the number of items in each part is not specified and k is even, the program sets $k_1 = k_2 = k/2$. If k is odd, $k_1 = (k+1)/2$. It is assumed that the first k_1 items are in part 1.

$W = \sum_{j=1}^{N} w_j$	Sum of the weights, where w_j is the weight for case j
$P_j = \sum_{i=1}^{k} X_{ji}$	The total score of the jth person
$\overline{P}_j = P_j / k$	Mean of the observations for the jth person
$T_i = \sum_{j=1}^{N} X_{ji} w_j$	The total score for the ith item
$G = \sum_{i=1}^{k} \sum_{j=1}^{N} X_{ji} w_j$	Grand sum of the scores
$\overline{G} = \dfrac{G}{Wk}$	Grand mean of the observations

Scale and Item Statistics—Method 1

Item Means and Standard Deviations

Mean for the *i*th Item

$$\overline{T}_i = T_i / W$$

Standard Deviation for the *i*th Item

$$S_i = \sqrt{\frac{\sum_{j=1}^{N} w_j X_{ji}^2 - W\overline{T}_i^2}{W-1}}$$

Scale Mean and Scale Variance

Scale Mean

$$M = G/W$$

For the split model:

Mean Part 1

$$M_1 = \sum_{i=1}^{k_1} \overline{T}_i$$

Mean Part 2

$$M_2 = \sum_{i=k_1+1}^{k} \overline{T}_i$$

Scale Variance

$$S_p^2 = \frac{1}{(W-1)}\left[\sum_{j=1}^{N} P_j^2 w_j - W\left(\sum_{i=1}^{k} \overline{T}_i\right)^2\right]$$

For the split model:

Variance Part 1

$$S_{p1}^2 = \frac{1}{W-1}\left[\sum_{j=1}^{N} w_j \left(\sum_{i=1}^{k_1} X_{ji}\right) - W\left(\sum_{i=1}^{k_1} \overline{T}_i\right)^2\right]$$

Variance Part 2

$$S_{p2}^2 = \frac{1}{W-1}\left[\sum_{j=1}^{N} w_j \left(\sum_{i=k_1+1}^{k} X_{ji}\right)^2 - W\left(\sum_{i=k_1+1}^{k} \overline{T}_i\right)^2\right]$$

Item-Total Statistics

Scale Mean if the *i*th Item is Deleted

$$\tilde{M}_i = M - \overline{T}_i$$

Scale Variance if the *i*th Item is Deleted

$$\tilde{S}_i^2 = S_p^2 + S_i^2 - 2\operatorname{cov}(X_i, P)$$

where the covariance between item *i* and the case score is

$$\operatorname{cov}(X_i, P) = \frac{1}{W-1}\left(\sum_{j=1}^{N} P_j X_{ji} w_j - \sum_{l=1}^{k} \overline{T}_l T_i\right)$$

Alpha if the *i*th Item Deleted

$$\overline{A}_i = \frac{k-1}{k-2}\left(1 - \sum_{\substack{l=1 \\ l \neq i}}^{k} S_l^2 \Big/ \tilde{S}_i^2\right)$$

Correlation between the *i*th Item and Sum of Others

$$R_i = \frac{\operatorname{cov}(X_i, P) - S_i^2}{S_i \tilde{S}_i}$$

The ANOVA Table (Winer, 1971)

Source of variation	Sum of Squares	df
Between people	$\sum_{j=1}^{N} P_j^2 w_j \Big/ k - G^2 \Big/ Wk$	$W-1$
Within people	$\sum_{i=1}^{k} \sum_{j=1}^{N} w_j X_{ji}^2 - \sum_{j=1}^{N} P_j^2 w_j \Big/ k$	$W(k-1)$
Between measures	$\sum_{i=1}^{k} T_i^2 \Big/ W - G^2 \Big/ Wk$	$k-1$
Residual	$\sum_{i=1}^{k} \sum_{j=1}^{N} w_j X_{ji}^2 - \sum_{j=1}^{N} P_j^2 w_j \Big/ k - \sum_{i=1}^{k} T_i^2 \Big/ W - G^2 \Big/ Wk$	$(W-1)(k-1)$
Total	$\sum_{i=1}^{k} \sum_{j=1}^{N} w_j X_{ji}^2 - G^2 \Big/ Wk$	$Wk-1$

Each of the mean squares is obtained by dividing the sum of squares by the corresponding degrees of freedom. The F ratio for between measures is

$$F = \frac{MS_{\text{between measures}}}{MS_{\text{residual}}}, \quad df = \left(k-1, (W-1)(k-1)\right)$$

Friedman Test or Cochran Test

$$\chi^2 = \frac{SS_{\text{between measures}}}{MS_{\text{within people}}}, \quad df = k-1$$

Note: Data must be ranks for the Friedman test and a dichotomy for the Cochran test.

Kendall's Coefficient of Concordance

$$W = \frac{SS_{\text{between measures}}}{SS_{\text{total}}}$$

(Will not be printed if Cochran is also specified.)

Tukey's Test for Nonadditivity

The residual sums of squares are further subdivided to

$$SS_{\text{nonadd}} = M^2 / D, \quad df = 1$$

where

$$D = \left(SS_{\text{bet. meas}} SS_{\text{bet. people}}\right) \big/ (Wk)$$

$$\left(= \left[\sum_{i=1}^{k} \left(\overline{T}_i - \overline{G}\right)^2 \right] \left[\sum_{j=1}^{N} w_j \left(\overline{P}_j - \overline{G}\right)^2 \right] \right)$$

$$M = \sum_{i=1}^{k} \overline{T}_i \sum_{j=1}^{N} \overline{P}_j X_{ji} w_j - \overline{G} \sum_{j=1}^{N} P_j^2 w_j \big/ k - \overline{G} SS_{\text{bet. meas}}$$

$$\left(= \sum_{j=1}^{N} w_j \left(\overline{P}_j - \overline{G}\right) \left[\sum_{i=1}^{k} X_{ji} \left(\overline{T}_i - \overline{G}\right) \right] \right)$$

$$SS_{\text{bal}} = SS_{\text{res}} - SS_{\text{nonadd}}, \quad df = (W-1)(k-1) - 1$$

The test for nonadditivity is

$$F = \frac{MS_{\text{nonadd}}}{MS_{\text{balance}}} \quad df = \left(1, (W-1)(k-1) - 1\right)$$

The regression coefficient for the nonadditivity term is

$$\hat{B} = M/D,$$

and the power to transform to additivity is

$$\hat{p} = 1 - \hat{B}\overline{G}$$

Scale Statistics

Reliability coefficient alpha (Cronbach 1951)

$$A = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} S_i^2}{S_p^2}\right)$$

If the model is split, separate alphas are computed:

$$A_1 = \frac{k_1}{k_1 - 1}\left(1 - \sum_{i=1}^{k_1} \frac{S_i^2}{S_{p1}^2}\right)$$

$$A_2 = \frac{k_2}{k_2 - 1}\left(1 - \sum_{i=k_1+1}^{k_2} \frac{S_i^2}{S_{p2}^2}\right)$$

For Split Model Only

Correlation Between the Two Parts of the Test

$$R = \frac{\frac{1}{2}\left(S_p^2 - S_{p1}^2 - S_{p2}^2\right)}{S_{p1} S_{p2}}$$

Equal Length Spearman-Brown Coefficient

$$Y = \frac{2R}{1+R}$$

Guttman Split Half

$$G = \frac{2\left(S_p^2 - S_{p1}^2 - S_{p2}^2\right)}{S_p^2}$$

Unequal Length Spearman-Brown

$$ULY = \frac{-R^2 + \sqrt{R^4 + 4R^2\left(1 - R^2\right)k_1 k_2 / k^2}}{2\left(1 - R^2\right)k_1 k_2 / k^2}$$

Basic Computations—Method 2

Items with zero variance are deleted from the scale and from k, k_1, and k_2. The inverses of matrices, when needed, are computed using the sweep operator described by Dempster (1969). If $|V| < 10^{-30}$, a warning is printed and statistics that require V^{-1} are skipped.

Covariance Matrix V and Correlation Matrix R

$$v_{ij} = \begin{cases} \left(\frac{1}{W-1}\left(\sum_{l=1}^{N} X_{li} X_{lj} w_j - W\overline{T}_i \overline{T}_j\right), i, j = 1, \ldots, k\right) & \text{if raw data input} \\ r_{ij} S_i S_j & \text{if correlation matrix and SD input} \end{cases}$$

$$r_{ij} = \frac{v_{ij}}{S_i S_i}, \quad \text{where} \quad S_i^2 = \frac{v_{ij}}{W-1}$$

Scale Variance

$$S_p^2 = \sum_{i=1}^{k} S_i^2 + 2\sum_{i<j}^{k}\sum^{k} v_{ij}$$

If the model is split,

$$S_{p1}^2 = \sum_{i=1}^{k_1} S_i^2 + 2\sum_{i<j}^{k_1}\sum^{k_1} v_{ij}$$

$$S_{p2}^2 = \sum_{i=k_1+1}^{k} S_i^2 + 2\sum_{i=k_1+1}^{k}\sum_{j>i}^{k} v_{ij}$$

where the first k_1 items are in part 1.

Scale Statistics—Method 2

Alpha Model

Estimated Reliability

$$\frac{k}{k-1}\left(1-\frac{\sum_{i=1}^{k} S_i^2}{S_p^2}\right)$$

Standardized Item Alpha

$$\frac{k\overline{\text{Corr}}}{1+(k-1)\overline{\text{Corr}}}$$

where

$$\overline{\text{Corr}} = \frac{2}{k(k-1)} \sum_{i<j}^{k} \sum^{k} r_{ij}$$

Split Model

Correlation between Forms

$$\frac{\sum_{i=1}^{k_1} \sum_{j=k_1+1}^{k} v_{ij}}{S_{p1} S_{p2}}$$

Guttman Split-Half

$$G = \frac{\sum_{i=1}^{k_1} \sum_{j=k_1+1}^{k} v_{ij}}{S_p^2}$$

Alpha and Spearman-Brown equal and unequal length are computed as in method 1.

Guttman Model (Guttman 1945)

$$L_1 = 1 - \frac{\sum_{i=1}^{k} S_i^2}{S_p^2}$$

$$L_2 = L_1 + \frac{\sqrt{\frac{2k}{k-1}\sum_{i<j}^{k}\sum^{k} v_{ij}^2}}{S_p^2}$$

$$L_3 = \frac{k}{k-1} L_1$$

$$L_4 = \frac{4\sum_{i<j}^{k}\sum^{k} v_{ij}}{S_p^2}$$

$$L_5 = L_1 + \frac{2\sqrt{\max_i \sum_{j \neq i}^{k} v_{ij}^2}}{S_p^2}$$

$$L_6 = 1 - \sum_{i=1}^{k} \varepsilon_i^2 \big/ S_p^2; \quad \text{where } \varepsilon_i^2 = \left(V^{-1}\right)_{ii}^{-1}$$

Parallel Model (Kristof 1963)

Common Variance

$$CV = \overline{\text{var}} = \frac{1}{k}\sum_{i=1}^{k} S_i^2$$

True Variance

$$TV = \overline{\text{cov}} = \frac{2}{k(k-1)} \sum^{k} \sum_{i<j}^{k} v_{ij}$$

Error Variance

$$EV = \overline{\text{var}} - \overline{\text{cov}}$$

Common Inter-Item Correlation

$$\hat{R} = \overline{\text{cov}} \big/ \overline{\text{var}}$$

Reliability of the Scale

$$A = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^{k} S_i^2}{S_p^2} \right)$$

Unbiased Estimate of the Reliability

$$\hat{A} = \frac{2 + (W-3)A}{(W-1)}$$

where A is defined above.

Test for Goodness of Fit

$$\chi^2 = -(W-1)\left(1 - \frac{k(k+1)^2(2k-3)}{12(k-1)\left(\frac{k(k+1)}{2}-2\right)(W-1)}\right)\log L$$

where

$$L = \frac{|V|}{\left(\overline{\text{var}} - \overline{\text{cov}}\right)^{k-1}\left(\overline{\text{var}} + (k+1)\overline{\text{cov}}\right)}$$

$$df = \frac{k(k+1)}{2} - 2$$

Log of the Determinant of the Unconstrained Matrix

$$\log UC = \log|V|$$

Log of the Determinant of the Constrained Matrix

$$\log C = \log\left(\left(\overline{\text{var}} - \overline{\text{cov}}\right)^{k-1}\left(\overline{\text{var}} + (k-1)\overline{\text{cov}}\right)\right)$$

Strict Parallel (Kristof 1963)

Common Variance

$$CV = \overline{\text{var}} + \frac{1}{k}\sum_{i=1}^{k}\left(\overline{T_i} - \overline{G}\right)^2$$

Error Variance

$$EV = MS_{\text{within people}}$$

All mean squares are calculated as described later in the analysis of variance table.

True Variance

$$TV = \overline{\text{var}} + \frac{1}{k}\sum_{i=1}^{k}\left(\overline{T_i} - \overline{G}\right)^2 - EV$$

Common Inter-Item Correlation

$$\hat{R} = \frac{\overline{\text{cov}} - \frac{1}{(k-1)k}\sum_{i=1}^{k}\left(\overline{T_i} - \overline{G}\right)^2}{\overline{\text{var}} + \frac{1}{k}\sum_{i=1}^{k}\left(\overline{T_i} - \overline{G}\right)^2}$$

Reliability of the Scale

$$\text{Re}\,l = \frac{k\hat{R}}{1 + (k-1)\hat{R}}$$

Unbiased Estimate of the Reliability

$$\text{Re}\,l = \frac{3 + (W-3)\,\text{Re}\,l}{W}$$

Test for Goodness of Fit

$$\chi^2 = -(W-1)\left(1-\frac{k(k+1)^2(2k-3)}{12(k-1)\left(k(k+3)/2-3\right)(W-1)}\right)\log L$$

where

$$L = \frac{|V|}{\left(\overline{\text{var}}+(k-1)\overline{\text{cov}}\right)\left(\overline{\text{var}}-\overline{\text{cov}}+\frac{1}{k}\sum_{i=1}^{k}\left(\overline{T}_i-\overline{G}\right)^2\right)^{k-1}}$$

$$df = k(k+3)/2-3$$

Log of the Determinant of the Unconstrained Matrix

$$\log UC = \log|V|$$

Log of the Determinant of the Constrained Matrix

$$\log C = \log\left(\overline{\text{var}}+(k-1)\overline{\text{cov}}\right)\left(\overline{\text{var}}-\overline{\text{cov}}+\frac{1}{k-1}\sum_{i=1}^{k}\left(\overline{T}_i-\overline{G}\right)^2\right)^{k-1}$$

Additional Statistics—Method 2

Descriptive and scale statistics and Tukey's test are calculated as in method 1. Multiple R^2 if an item is deleted is calculated as

$$\tilde{R}_i^2 = 1 - \frac{\varepsilon_i^2}{S_i^2} \qquad \varepsilon_i^2 = \frac{1}{\left(V^{-1}\right)_{ii}}$$

Analysis of Variance Table

Source of variation	Sum of Squares	df
Between people	$(W-1)\left[\frac{1}{k}\left(\sum\sum_{i=1}^{k} S_i^2 - \frac{2}{k-1}\sum\sum_{i<j} v_{ij}\right)\right] + \frac{2}{(k-1)}\sum_{i<j} v_{ij}$	$W-1$
Within people	$\frac{(W-1)(k-1)}{k}\left[\sum_{i=1}^{k} S_i^2 - \frac{2}{k-1}\sum\sum_{i<j} v_{ij}\right] + (W-1)SS_{\text{bet. people}}$	$W(k-1)$
Between measures	$W\left(\sum_{i=1}^{k} \overline{T}_i^2 - \frac{1}{k}\left(\sum_{i=1}^{k} \overline{T}_i\right)^2\right)$	$k-1$
Residual	$\frac{(W-1)(k-1)}{k}\left[\sum_{i=1}^{k} S_i^2 - \frac{2}{k-1}\sum\sum_{i<j} v_{ij}\right]$	$(W-1)(k-1)$
Total	Between *SS* + Within *SS*	$Wk-1$

Hotelling's T^2 (Winer, 1971)

$$T^2 = W\mathbf{Y}'\mathbf{B}^{-1}\mathbf{Y}$$

where

$$\mathbf{Y} = \begin{bmatrix} \overline{T}_1 - \overline{T}_k \\ \overline{T}_2 - \overline{T}_k \\ \vdots \\ \overline{T}_{k-1} - \overline{T}_k \end{bmatrix}$$

$$\mathbf{B} = \mathbf{CVC}'$$

where **C** is an identity matrix of rank $k-1$ augmented with a column of –1 on the right.

$$b_{ij} = v_{ij} - v_{ik} - v_{jk} + S_k^2$$

The test will not be done if $W < k$ or $|\mathbf{B}| < 10^{-30}$.

The significance level of T^2 is based on

$$F = \frac{W-k+1}{(W-1)(k-1)} T^2, \text{ with } df = (k-1, W-k+1)$$

Item Mean Summaries

$$\text{Mean} = \sum_{i=1}^{k} \overline{T}_i \Big/ k$$

$$\text{Variance} = \frac{\sum_{i=1}^{k} \overline{T}_i^{\,2} - \left(\sum_{i=1}^{k} \overline{T}_i \right)^2 \Big/ k}{(k-1)}$$

$$\text{Maximum} = \max_i \overline{T}_i$$

$$\text{Minimum} = \min_i \overline{T}_i$$

$$\text{Range} = \text{Maximum} - \text{Minimum}$$

$$\text{Ratio} = \frac{\text{Maximum}}{\text{Minimum}}$$

Item Variance Summaries

Same as for item means excepts that S_i^2 is substituted for $\overline{T}_i$ in all calculations.

Inter-Item Covariance Summaries

$$\text{Mean} = \frac{\sum\sum_{i<j} v_{ij}}{k(k-1)}$$

$$\text{Variance} = \frac{1}{k(k-1)-1}\left[\sum\sum_{i<j} v_{ij}^2 - \frac{1}{k(k-1)}\left(\sum\sum_{i<j} v_{ij}\right)^2\right]$$

$$\text{Maximum} = \max_{i,j} v_{ij}$$

$$\text{Minimum} = \min_{i,j} v_{ij}$$

$$\text{Range} = \text{Maximum} - \text{Minimum}$$

$$\text{Ratio} = \frac{\text{Maximum}}{\text{Minimum}}$$

Inter-Item Correlations

Same as for inter-item covariances, with v_{ij} being replaced by r_{ij}.

If the model is split, statistics are also calculated separately for each scale.

Intraclass Correlation Coefficients

Intraclass correlation coefficients are always discussed in a random/mixed effects model setting. McGraw and Wong (1996) is the key reference for this document. See also Shrout and Fleiss (1979).

In this document, two measures of correlation are given for each type under each model: **single measure** and **average measure**. Single measure applies to single measurements, for example, the ratings of judges, individual item scores, or the body weights of individuals, whereas average measure applies to average measurements, for example, the average rating for k judges, or the average score for a k-item test.

One-Way Random Effects Model: People Effect Random

Let X_{ji} be the response to the i-th measure given by the j-th person, $i = 1, \ldots, k$, $j = 1, \ldots, W$. Suppose that X_{ji} can be expressed as $X_{ji} = \mu + p_j + w_{ji}$, where p_j is the between-people effect which is normal distributed with zero mean and a variance of σ_p^2, and w_{ji} is the within-people effect which is also normal distributed with zero mean and a variance of σ_w^2.

Let MS_{BP} and MS_{WP} be the respective between-people Mean Squares and within-people Mean Squares. Formulas for these two quantities can be found on page 479 of *SPSS 7.5 Statistical Algorithms* by dividing the corresponding Sum of Squares with its degrees of freedom.

Single Measure Intraclass Correlation

The single measure intraclass correlation is defined as

$$\rho_{(1)} = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_w^2}.$$

Estimate

The single measure intraclass correlation coefficient is estimated by

$$ICC(1) = \frac{MS_{\mathrm{BP}} - MS_{\mathrm{WP}}}{MS_{\mathrm{BP}} + (k-1)\,MS_{\mathrm{WP}}}.$$

In general,

$$\frac{-1}{k-1} < ICC(1) \le 1.$$

Confidence Interval

For 0 < α < 1, a (1-α)100% confidence interval for $\rho_{(1)}$ is given by

$$\frac{F_{p/w} - F_{\alpha/2,W-1,W(k-1)}}{F_{p/w} + (k-1)F_{\alpha/2,W-1,W(k-1)}} < \rho_{(1)} < \frac{F_{p/w} - F_{1-\alpha/2,W-1,W(k-1)}}{F_{p/w} + (k-1)F_{1-\alpha/2,W-1,W(k-1)}},$$

where

$$F_{p/w} = \frac{MS_{\mathrm{BP}}}{MS_{\mathrm{WP}}}$$

and F_{α',v_1,v_2} is the upper α' point of a F-distribution with degrees of freedom v_1 and v_2.

Hypothesis Testing

The test statistic $F^{(1)}$ for $H_0{:}\,\rho_{(1)} = \rho_0$, where $1 > \rho_0 \geq 0$ is the hypothesized value, is

$$F^{(1)} = F_{p/w}\frac{1-\rho_0}{1+(k-1)\rho_0}.$$

Under the null hypothesis, the test statistic has an F-distribution with $W-1, W(k-1)$ degrees of freedom.

Average Measure Intraclass Correlation

The average measure intraclass correlation is defined as

$$\rho_{(k)} = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_w^2 / k}.$$

Estimate

The average measure intraclass correlation coefficient is estimated by

$$ICC(k) = \frac{MS_{\text{BP}} - MS_{\text{WP}}}{MS_{\text{BP}}}.$$

Confidence Interval

A (1-α)100% confidence interval for $\rho_{(k)}$ is given by

$$\frac{F_{p/w} - F_{\alpha/2,W-1,W(k-1)}}{F_{p/w}} < \rho_{(k)} < \frac{F_{p/w} - F_{1-\alpha/2,W-1,W(k-1)}}{F_{p/w}}.$$

Hypothesis Testing

The test statistic $F^{(k)}$ for $H_0{:}\rho_{(k)} = \rho_0$, where $1 > \rho_0 \ge 0$ is the hypothesized value, is

$$F^{(k)} = F_{p/w}(1 - \rho_0).$$

Under the null hypothesis, the test statistic has an *F*-distribution with $W-1, W(k-1)$ degrees of freedom.

Two-Way Random Effects Model: People and Measures Effects Random

Let X_{ji} be the response to the *i*-th measure given by the *j*-th person, i = 1, …, k, j = 1, …, W. Suppose that X_{ji} can be expressed as $X_{ji} = \mu + p_j + m_i + pm_{ji} + e_{ji}$, where p_j is the *people effect* which is normal distributed with zero mean and a variance of σ_p^2, m_i is the *measures effect* which is normal distributed with zero mean and a variance of σ_m^2, pm_{ji} is the interaction effect which is normal distributed with zero mean and a variance of σ_{pm}^2, and e_{ji} is the error effect which is again normal distributed with zero mean and a variance of σ_e^2.

Let MS_{BP}, MS_{BM} and MS_{Res} be the respective between-people Mean Squares, between-measures Mean Squares and Residual Mean Squares. Formulas for these quantities can be found on page 479 of *SPSS 7.5 Statistical Algorithms* by dividing the corresponding Sum of Squares with its degrees of freedom.

Type A Single Measure Intraclass Correlation

The type A single measure intraclass correlation is defined as

$$\rho_{(A,1,r)} = \begin{cases} \dfrac{\sigma_p^2}{\sigma_p^2 + \sigma_m^2 + \sigma_{pm}^2 + \sigma_e^2} & \text{if interaction effect } pm_{ji} \text{ is present} \\ \dfrac{\sigma_p^2}{\sigma_p^2 + \sigma_m^2 + \sigma_e^2} & \text{if interaction effect } pm_{ji} \text{ is absent} \end{cases}.$$

Estimate

The type A single measure intraclass correlation coefficient is estimated by

$$ICC(A,1,r) = \frac{MS_{\mathrm{BP}} - MS_{\mathrm{Res}}}{MS_{\mathrm{BP}} + (k-1)MS_{\mathrm{Res}} + k(MS_{\mathrm{BM}} - MS_{\mathrm{Res}})/W}.$$

Notice that the same estimator is used whether or not the interaction effect pm_{ji} is present.

Confidence Interval

A (1-α)100% confidence interval is given by

$$\frac{W(MS_{\mathrm{BP}} - F_{\alpha/2,W-1,v} \cdot MS_{\mathrm{Res}})}{F_{\alpha/2,W-1,v}[k \cdot MS_{\mathrm{BM}} + (kW-k-W)MS_{\mathrm{Res}}] + W \cdot MS_{\mathrm{BP}}} < \rho_{(A,1,r)} < \frac{W(MS_{\mathrm{BP}} - F_{1-\alpha/2,W-1,v} \cdot MS_{\mathrm{Res}})}{F_{1-\alpha/2,W-1,v}[k \cdot MS_{\mathrm{BM}} + (kW-k-W)MS_{\mathrm{Res}}] + W \cdot MS_{\mathrm{BP}}},$$

where

$$v = \frac{(aMS_{\mathrm{BM}} + bMS_{\mathrm{Res}})^2}{\left[\dfrac{(aMS_{\mathrm{BM}})^2}{k-1} + \dfrac{(bMS_{\mathrm{Res}})^2}{(W-1)(k-1)}\right]}$$

$$a = \frac{k \cdot ICC(A,1,r)}{W(1 - ICC(A,1,r))}$$

and

$$b = 1 + \frac{k \cdot ICC(A,1,r) \cdot (W-1)}{W(1 - ICC(A,1,r))}.$$

Hypothesis Testing

The test statistic $F^{(A,1,r)}$ for $H_0 : \rho_{(A,1,r)} = \rho_0$, where $1 > \rho_0 \geq 0$ is the hypothesized value, is

$$F^{(A,1,r)} = \frac{MS_{\mathrm{BP}}}{a_0 MS_{\mathrm{BM}} + b_0 MS_{\mathrm{Res}}}$$

where

$$a_0 = \frac{k\rho_0}{W(1-\rho_0)}$$

and

$$b_0 = 1 + \frac{k\rho_0 (W-1)}{W(1-\rho_0)}.$$

Under the null hypothesis, the test statistic has an *F*-distribution with $W-1, \nu_0^{(1)}$ degrees of freedom.

$$\nu_0^{(1)} = \frac{(a_0 MS_{\mathrm{BM}} + b_0 MS_{\mathrm{Res}})^2}{\left[\dfrac{(a_0 MS_{\mathrm{BM}})^2}{k-1} + \dfrac{(b_0 MS_{\mathrm{Res}})^2}{(W-1)(k-1)}\right]}.$$

Type A Average Measure Intraclass Correlation

The type A average measure intraclass correlation is defined as

$$\rho_{(A,k,r)} = \begin{cases} \dfrac{\sigma_p^2}{\sigma_p^2 + (\sigma_m^2 + \sigma_{pm}^2 + \sigma_e^2)/k} & \text{if interaction effect } pm_{ji} \text{ is present} \\ \dfrac{\sigma_p^2}{\sigma_p^2 + (\sigma_m^2 + \sigma_e^2)/k} & \text{if interaction effect } pm_{ji} \text{ is absent} \end{cases}.$$

Estimate

The type A average measure intraclass correlation coefficient is estimated by

$$ICC(A,k,r) = \frac{MS_{\text{BP}} - MS_{\text{Res}}}{MS_{\text{BP}} + (MS_{\text{BM}} - MS_{\text{Res}})/W}.$$

Notice that the same estimator is used whether or not the interaction effect pm_{ji} is present.

Confidence Interval

A (1- α)100% confidence interval is given by

$$\frac{W(MS_{\text{BP}} - F_{\alpha/2,W-1,v}MS_{\text{Res}})}{F_{\alpha/2,W-1,v}(MS_{\text{BM}} - MS_{\text{Res}}) + W \cdot MS_{\text{BP}}} < \rho_{(A,k,r)} < \frac{W(MS_{\text{BP}} - F_{1-\alpha/2,W-1,v}MS_{\text{Res}})}{F_{1-\alpha/2,W-1,v}(MS_{\text{BM}} - MS_{\text{Res}}) + W \cdot MS_{\text{BP}}}$$

where v is defined as in the Type A single measure confidence interval, with $ICC(A,1,r)$ replaced by $ICC(A,k,r)$.

Hypothesis Testing

The test statistic for $H_0{:}\rho_{(A,k,r)} = \rho_0$, where $1 > \rho_0 \geq 0$ is the hypothesized value, is

$$F^{(A,k,r)} = \frac{MS_{\text{BP}}}{c_0 MS_{\text{BM}} + d_0 MS_{\text{Res}}}$$

where

$$c_0 = \frac{\rho_0}{W(1-\rho_0)}$$

and

$$d_0 = 1 + \frac{\rho_0 (W-1)}{W(1-\rho_0)}.$$

Under the null hypothesis, the test statistic has an F-distribution with $W-1, v_0^{(k)}$ degrees of freedom.

$$v_0^{(k)} = \frac{(c_0 MS_{\text{BM}} + d_0 MS_{\text{Res}})^2}{\left[\frac{(c_0 MS_{\text{BM}})^2}{k-1} + \frac{(d_0 MS_{\text{Res}})^2}{(W-1)(k-1)}\right]}.$$

Type C Single Measure Intraclass Correlation

The type C single measure intraclass correlation is defined as

$$\rho_{(C,1,r)} = \begin{cases} \dfrac{\sigma_p^2}{\sigma_p^2 + \sigma_{pm}^2 + \sigma_e^2} & \text{if interaction effect } pm_{ji} \text{ is present} \\ \dfrac{\sigma_p^2}{\sigma_p^2 + \sigma_e^2} & \text{if interaction effect } pm_{ji} \text{ is absent} \end{cases}.$$

Estimate

The type C single measure intraclass correlation coefficient is estimated by

$$ICC(C,1,r) = \frac{MS_{\mathrm{BP}} - MS_{\mathrm{Res}}}{MS_{\mathrm{BP}} + (k-1)\,MS_{\mathrm{Res}}}.$$

Notice that the same estimator is used whether or not the interaction effect pm_{ji} is present.

Confidence Interval

A (1-α)100% confidence interval is given by

$$\frac{F_{p/r} - F_{\alpha/2,W-1,(W-1)(k-1)}}{F_{p/r} + (k-1)F_{\alpha/2,W-1,(W-1)(k-1)}} < \rho_{(C,1,r)} < \frac{F_{p/r} - F_{1-\alpha/2,W-1,(W-1)(k-1)}}{F_{p/r} + (k-1)F_{1-\alpha/2,W-1,(W-1)(k-1)}}$$

where

$$F_{p/r} = \frac{MS_{\mathrm{BP}}}{MS_{\mathrm{Res}}}.$$

Hypothesis Testing

The test statistic for $H_0 : \rho_{(C,1,r)} = \rho_0$, where $1 > \rho_0 \geq 0$ is the hypothesized value, is

$$F^{(C,1,r)} = F_{p/r}\,\frac{1-\rho_0}{1+(k-1)\rho_0}.$$

Under the null hypothesis, the test statistic has an *F*-distribution with $W-1, (W-1)(k-1)$ degrees of freedom.

Type C Average Measure Intraclass Correlation

The type C average measure intraclass correlation is defined as

$$\rho_{(C,k,r)} = \begin{cases} \dfrac{\sigma_p^2}{\sigma_p^2 + (\sigma_{pm}^2 + \sigma_e^2)/k} & \text{if interaction effect } pm_{ji} \text{ is present} \\ \dfrac{\sigma_p^2}{\sigma_p^2 + \sigma_e^2/k} & \text{if interaction effect } pm_{ji} \text{ is absent} \end{cases}.$$

Estimate

The type C average measure intraclass correlation coefficient is estimated by

$$ICC(C,k,r) = \frac{MS_{\text{BP}} - MS_{\text{Res}}}{MS_{\text{BP}}}.$$

Notice that the same estimator is used whether or not the interaction effect pm_{ji} is present.

Confidence Interval

A (1- α)100% confidence interval is given by

$$\frac{F_{p/r} - F_{\alpha/2,W-1,(W-1)(k-1)}}{F_{p/r}} < \rho_{(C,k,r)} < \frac{F_{p/r} - F_{1-\alpha/2,W-1,(W-1)(k-1)}}{F_{p/r}}.$$

Hypothesis Testing

The test statistic for $H_0: \rho_{(C,k,r)} = \rho_0$, where $1 > \rho_0 \geq 0$ is the hypothesized value, is

$$F^{(C,k,r)} = F_{p/r}(1-\rho_0).$$

Under the null hypothesis, the test statistic has an F-distribution with $W-1,(W-1)(k-1)$ degrees of freedom.

Two-Way Mixed Effects Model: People Effects Random, Measures Effects Fixed

Let X_{ji} be the response to the i-th measure given by the j-th person, i = 1, …, k, j = 1, …, W. Suppose that X_{ji} can be expressed as $X_{ji} = \mu + p_j + m_i + pm_{ji} + e_{ji}$, where p_j is the *people effect* which is normal distributed with zero mean and a variance of σ_p^2, m_i is considered as a fixed effect, pm_{ji} is the interaction effect which is normal distributed with zero mean and a variance of σ_{pm}^2, and e_{ji} is the error effect which is again normal distributed with zero mean and a variance of σ_e^2. Denote θ_m^2 as the expected measure square of between measures effect m_i.

Let MS_{BP} and MS_{Res} be the respective between-people Mean Squares and Residual Mean Squares. Formulas for these quantities can be found on page 479 of *SPSS 7.5 Statistical Algorithms* by dividing the corresponding Sum of Squares with its degrees of freedom.

Type A Single Measure Intraclass correlation

The type A single measure Intraclass correlation is defined as

$$\rho_{(A,1,m)} = \begin{cases} \dfrac{\sigma_p^2 - \sigma_{pm}^2/(k-1)}{\sigma_p^2 + \theta_m^2 + (\sigma_{pm}^2 + \sigma_e^2)} & \text{if interaction effect } pm_{ji} \text{ is present} \\[2ex] \dfrac{\sigma_p^2}{\sigma_p^2 + \theta_m^2 + \sigma_e^2} & \text{if interaction effect } pm_{ji} \text{ is absent} \end{cases}.$$

Estimate

The type A single measure intraclass correlation is estimated by

$$ICC(A,1,m) = \frac{MS_{\text{BP}} - MS_{\text{Res}}}{MS_{\text{BP}} + (k-1)MS_{\text{Res}} + k(MS_{\text{BM}} - MS_{\text{Res}})/W}.$$

Notice that the same estimator is used whether or not the interaction effect pm_{ji} is present.

Confidence Interval

A (1-α)100% confidence interval for $\rho_{(A,1,m)}$ is the same as that for $\rho_{(A,1,r)}$, with $ICC(A,1,r)$ replaced by $ICC(A,1,m)$.

Hypothesis Testing

The test statistic for $H_0{:}\rho_{(A,1,m)} = \rho_0$, where $1 > \rho_0 \geq 0$ is the hypothesized value, is the same as that for $\rho_{(A,1,r)}$, with the same distribution under the null hypothesis.

Type A Average Measure Intraclass Correlation

The type A average measure intraclass correlation is defined as

$$\rho_{(A,k,m)} = \begin{cases} \dfrac{\sigma_p^2 - \sigma_{pm}^2/(k-1)}{\sigma_p^2 + (\theta_m^2 + \sigma_{pm}^2 + \sigma_e^2)/k} & \text{if interaction effect } pm_{ji} \text{ is present} \\ \dfrac{\sigma_p^2}{\sigma_p^2 + (\theta_m^2 + \sigma_e^2)/k} & \text{if interaction effect } pm_{ji} \text{ is absent} \end{cases}.$$

Estimate

The type A single measure intraclass correlation is estimated by

$$ICC(A,k,m) = \begin{cases} \text{Not estimable} & \text{if interaction effect } pm_{ji} \text{ is present} \\ \dfrac{MS_{\text{BP}} - MS_{\text{Res}}}{MS_{\text{BP}} + (MS_{\text{BM}} - MS_{\text{Res}})/W} & \text{if interaction effect } pm_{ji} \text{ is absent} \end{cases}$$

.

Confidence Interval

A (1-α)100% confidence interval for $\rho_{(A,k,m)}$ is the same as that for $\rho_{(A,k,r)}$, with $ICC(A,k,r)$ replaced by $ICC(A,k,m)$. Notice that the hypothesis test is not available when the interaction effect pm_{ji} is present.

Hypothesis Testing

The test statistic for $H_0{:}\rho_{(A,k,m)} = \rho_0$, where $1 > \rho_0 \geq 0$ is the hypothesized value, is the same as that for $\rho_{(A,k,r)}$, with the same distribution under the null hypothesis. Notice that the hypothesis test is not available when the interaction effect pm_{ji} is present.

Type C Single Measure Intraclass Correlation

The type C single measure intraclass correlation is defined as

$$\rho_{(C,1,m)} = \begin{cases} \dfrac{\sigma_p^2 - \sigma_{pm}^2 / (k-1)}{\sigma_p^2 + (\sigma_{pm}^2 + \sigma_e^2)} & \text{if interaction effect } pm_{ji} \text{ is present} \\ \dfrac{\sigma_p^2}{\sigma_p^2 + \sigma_e^2} & \text{if interaction effect } pm_{ji} \text{ is absent} \end{cases}.$$

Estimate

The type C single measure intraclass correlation is estimated by

$$ICC(C,1,m) = \frac{MS_{\text{Between people}} - MS_{\text{Residual}}}{MS_{\text{Between people}} + (k-1)\,MS_{\text{Residual}}}.$$

Notice that the same estimator is used whether or not the interaction effect pm_{ji} is present.

Confidence Interval

A (1-α)100% confidence interval is given by

$$\frac{F_{p/r} - F_{\alpha/2,W-1,(W-1)(k-1)}}{F_{p/r} + (k-1)F_{\alpha/2,W-1,(W-1)(k-1)}} < \rho_{(C,1,m)} < \frac{F_{p/r} - F_{1-\alpha/2,W-1,(W-1)(k-1)}}{F_{p/r} + (k-1)F_{1-\alpha/2,W-1,(W-1)(k-1)}}.$$

where

$$F_{p/r} = \frac{MS_{\text{BP}}}{MS_{\text{Res}}}.$$

Hypothesis Testing

The test statistic for $H_0 : \rho_{(C,1,m)} = \rho_0$, where $1 > \rho_0 \geq 0$ is the hypothesized value, is

$$F^{(C,1,m)} = F_{p/r} \frac{1-\rho_0}{1+(k-1)\rho_0}.$$

Under the null hypothesis, the test statistic has an F-distribution with $W-1, (W-1)(k-1)$ degrees of freedom.

Type C Average Measure Intraclass Correlation

The type C average measure intraclass correlation is defined as

$$\rho_{(C,k,m)} = \begin{cases} \dfrac{\sigma_p^2 - \sigma_{pm}^2/(k-1)}{\sigma_p^2 + (\sigma_{pm}^2 + \sigma_e^2)/k} & \text{if interaction effect } pm_{ji} \text{ is present} \\ \dfrac{\sigma_p^2}{\sigma_p^2 + \sigma_e^2/k} & \text{if interaction effect } pm_{ji} \text{ is absent} \end{cases}.$$

Estimate

The type C average measure intraclass correlation is estimated by

$$ICC(C,k,m) = \begin{cases} \text{Not estimable} & \text{if interaction effect } pm_{ji} \text{ is present} \\ \dfrac{MS_{\text{BP}} - MS_{\text{Res}}}{MS_{\text{BP}}} & \text{if interaction effect } pm_{ji} \text{ is absent} \end{cases}.$$

Confidence Interval

A (1-α)100% confidence interval is given by

$$\frac{F_{p/r} - F_{\alpha/2, W-1, (W-1)(k-1)}}{F_{p/r}} < \rho_{(C,k,m)} < \frac{F_{p/r} - F_{1-\alpha/2, W-1, (W-1)(k-1)}}{F_{p/r}}.$$

Notice that the confidence interval is not available when the interaction effect pm_{ji} is present.

Hypothesis Testing

The test statistic for $H_0: \rho_{(C,k,m)} = \rho_0$, where $1 > \rho_0 \geq 0$ is the hypothesized value, is

$$F^{(C,1,m)} = F_{p/r}(1 - \rho_0).$$

Under the null hypothesis, the test statistic has an F-distribution with $W-1, (W-1)(k-1)$ degrees of freedom. Notice that the F-test is not available when the interaction effect pm_{ji} is present.

References

Cronbach, L. J. 1951. Coefficient alpha and internal structure of tests. *Psychometrika*, 16: 297–334.

Dempster, A. P. 1969. *Elements of Continuous Multivariate Analysis*. Reading, Mass.: Addison-Wesley.

Guttman, L. 1945. A basis for analysing test-retest reliability. *Psychometrika*, **10**(4): 255–282.

Kristof, W. 1963. The statistical theory of stepped-up reliability coefficients when the test has been divided into several equivalent parts. *Psychometrika*, 28: 221–238.

Kristof, W. 1969. Estimation of true score and error variance under various equivalent assumptions. *Psychometrika*, 34: 489–507.

McGraw, K. O., and Wong, S. P. (1996). Forming Inferences About Some Intraclass Correlation Coefficients. *Psychological Methods*, 1, p30-46.

Novick, M. R., and Lewis, C. 1967. Coefficient alpha and the reliability of composite measurements. *Psychometrika*, 32: 1–13.

Shrout, P. E., and Fleiss, J. L. (1979). Intraclass Correlations: Uses in Assessing Rater Reliability. *Psychological Bulletin*, 86, p420-428.

Winer, B. J. 1971. *Statistical principles in experimental design*, 2nd ed. New York: McGraw-Hill.