

SPSS programs for the measurement of nonindependence in standard dyadic designs

VALENTIM R. ALFERES

University of Coimbra, Coimbra, Portugal

AND

DAVID A. KENNY

University of Connecticut, Storrs, Connecticut

Dyadic research is becoming more common in the social and behavioral sciences. The most common dyadic design is one in which two persons are measured on the same set of variables. Very often, the first analysis of dyadic data is to determine the extent to which the responses of the two persons are correlated—that is, whether there is nonindependence in the data. We describe two user-friendly SPSS programs for measuring nonindependence of dyadic data. Both programs can be used for distinguishable and indistinguishable dyad members. *Inter1.sps* is appropriate for interval measures. *Inter2.sps* applies to categorical variables. The SPSS syntax and data files related to this article may be downloaded as supplemental materials from brm.psychonomic-journals.org/content/supplemental.

Dyadic research is becoming more common in the social and behavioral sciences. Researchers study persons who pair themselves (e.g., married couples, friends), laboratory dyads (e.g., two people involved in a negotiation), two persons who do not interact but receive the same experimental information (e.g., yoked pairs), and even two aspects of the same person (e.g., recordings from both hemispheres of the brain or from both eyes).

Dyads are said to be distinguishable when the two persons from all the dyads can be ordered in the same way. For instance, for heterosexual married persons, the data can be ordered by gender. Dyads are said to be indistinguishable when the two persons from all the dyads cannot be ordered in the same way. For instance, roommates and friends are usually treated as indistinguishable.

According to Kenny, Kashy, and Cook (2006), the most common dyadic design is one in which each person is paired to one other person and these two persons are measured on the same set of variables. The first analysis of dyadic data is to determine the extent to which the responses of the two persons are correlated—that is, whether there is nonindependence in the data. Information about nonindependence has important theoretical and statistical implications. Theoretically, nonindependence can be used to infer reciprocity, synchrony, or influence in the dyad. Statistically, nonindependence requires that the data be analyzed in ways that include both dyad and person as units of analysis.

SPSS Programs for the Measurement of Nonindependence

On the basis of the systematization of nonindependence measures presented in Kenny et al. (2006), two user-

friendly SPSS programs for preliminary analysis of reciprocal standard dyadic designs were created. *Inter1.sps* is appropriate for interval-level measures. *Inter2.sps* handles categorical variables. Both programs can be used for distinguishable and indistinguishable dyad members.

Inter1.sps: Interval measures. *Inter1.sps* is an SPSS syntax file that runs automatically from the SPSS syntax window. The only prerequisite for the user is to have an active SPSS file containing three variable columns in dyadic data format, named *dyad*, *x.1*, and *x.2* (see the data matrix in Appendix A). Note that each record refers to a dyad and not to an individual. The file may contain other variables in the data set. *Dyad* is an identifier variable for the dyad (dyads 1 to *n*); *x.1* and *x.2* are the scores for Member 1 and Member 2 on the focus variable, measured on an interval-level scale. There must be no missing data for *dyad*, *x.1*, or *x.2*.¹

For dyads with distinguishable members, the program computes the standard Pearson product-moment correlation coefficient (*r*) between dyad members' scores and performs a *t* test of the null hypothesis that the population correlation is 0. Additionally, *Inter1.sps* computes lower and upper bounds of the correlation coefficient confidence interval, using the *r* to Fisher's *z* transformation.

For dyads with indistinguishable members, the program uses the two approaches detailed in Kenny et al. (2006). In the first approach, nonindependence is assessed by the intraclass correlation coefficient (ICC), which is computed in the context of the ANOVA one-way random-effects model. The resulting coefficient is ICC(1,1), according to the terminology introduced by Shrout and Fleiss (1979). The same authors give the formulas used in significance testing and

confidence interval estimation. Note that before calculating ICC, the program changes the file layout to an individual data structure (for an illustration of individual, dyadic, and pairwise data structures, see Kenny et al., 2006).

An alternative approach to the measurement of non-independence when members of a dyad are indistinguishable is based on the double-entry method and is known as *pairwise correlation* (r_p ; Gonzalez & Griffin, 1999; Griffin & Gonzalez, 1995). More precisely, the initial dyadic structure is vertically duplicated, changing Member 1 (x.1) and Member 2 (x.2) scores in the second half of the data matrix, and r is computed on pairwise data ($2n$, instead of n , pairs of scores). Inter1.sps automatically changes the initial file dyadic layout to a pairwise data structure, computes r_p , and adopts the Griffin and Gonzalez approach to adjust significance testing (z test) and confidence interval estimates for n pairs of scores.

By default, α is set to the conventional value of .05. However, the user can easily modify α for confidence intervals in the corresponding command line at the beginning of the program.

Although r and the corresponding p value could have been obtained directly from the menus available in SPSS Data Editor—namely, the Analyze/Correlate/Bivariate menu—the user cannot get the confidence interval for r from SPSS graphic interface. When members of a dyad are indistinguishable, ICC is computed in the Analyze/Scale/Reliability Analysis menu if the data file has a dyadic structure and the user specifies the correct options (Model = Alpha | Statistics/Intraclass correlation coefficient/Model = One-Way Random). However, SPSS provides a one-tailed p value instead of the correct two-tailed probability (Kenny et al., 2006) and treats large negative values as not significant. Moreover, it can be problematic to use SPSS to compute r_p via the Analyze/Correlate/Bivariate menu, because the significance test has to be adjusted for n and not $2n$ dyads. The advantage of Inter1.sps is that this adjustment is made and the confidence interval bounds are properly calculated.

Inter2.sps: Categorical measures. When the level of measurement for the variable of interest is nominal (e.g., both members of a dyad respond to a question about their voting behavior in the last election), assessment of non-independence is done by quantifying agreement between members and correcting this quantity for expected random agreements. The measure of agreement for categorical variables is Cohen's κ (Cohen, 1960). Inter2.sps is an SPSS syntax file that computes κ in situations of both distinguishable and indistinguishable members.

For dyads with distinguishable members, computing and testing κ is straightforward. If the data are raw data, the user can obtain κ and the corresponding p value directly from the SPSS graphic interface (Analyze/Descriptive Statistics/Crosstabs/Statistics/Kappa). Alternatively, if the data are already summarized as a cross-table between Member 1 and Member 2 response categories (see Appendix B for SPSS Data Editor Layout and Appendix C, Table C2, for the corresponding cross-table), the same output can be reached if the user weights the variables by the cells' frequencies before running the analysis.

The analysis in the indistinguishable-members situation is a little more complicated, because the assignment of Member 1 and Member 2 response categories to the rows and columns of a cross-table is essentially arbitrary. Kenny et al. (2006) suggest that κ can be computed on a symmetrical matrix, obtained by averaging the off-diagonal cells (nonagreements) of the original cross-table (see the two cross-tabulations in Appendix C, Table C2). After this rearrangement, calculation of κ is identical to that in the distinguishable-members situation.

Inter2.sps computes κ in the same way it is computed in the SPSS graphic interface, with the exception that the off-diagonal averaging is done automatically. The input data for Inter2.sps must be a file with three variables named *member1*, *member2* and *count* (see the data matrix in Appendix B). The first two variables are the members' response categories; the third variable corresponds to cell frequencies in the standard cross-table (all cell frequencies must be greater than 0). To obtain this table from nontabulated raw data in order to prepare the input file, the Analyze/Descriptive Statistics/Crosstabs menu is available from the SPSS graphic interface.²

The output gives κ , and the corresponding statistical significance test ([Approx. T] is a one-tailed z test; the standard error under the null hypothesis is $s_{\kappa n} = [\text{Value}]/[\text{Approx. T}]$). To compute the confidence interval of κ , the user must take the standard error not assuming the null hypothesis from the output ($s_k = [\text{Asymp. Std. Error}]$). For example, the 95% interval confidence bounds of κ are $\kappa \pm 1.96s_k$.

Programs testing and availability. Inter1.sps and Inter2.sps were successfully tested with Versions 14.0.2, 15.0.1, and 16.0.2 of SPSS for Windows. The programs, as well as complementary documentation, can be downloaded from the authors' Web sites at <http://davidakenny.net/kkc/c2/c2.htm> or www.fpce.uc.pt/niips/inter.htm.

Illustrative Examples

Inter1.sps and Inter2.sps are briefly illustrated with short and hypothetical data sets, presented in Kenny et al. (2006). The first data set, used with Inter1.sps, is taken from Table 2.1. Suppose that the data are mutual liking scores on a 9-point scale. In the distinguishable-members situation, the data came from 10 romantic heterosexual couples. In the indistinguishable-members situation, the same data were collected from 10 pairs of same-sex roommates. The dyadic data are reproduced in Appendix A; selected output from Inter1.sps, including r , ICC, and r_p , is shown in Appendix C, Table C1.

The second data set, used with Inter2.sps, is taken from Table 2.4 (Kenny et al., 2006) and represents 200 dyads' hypothetical responses on a nominal variable with three categories (A, B, and C). For the distinguishable- and indistinguishable-members situations, suppose that the dyads are, again and respectively, romantic heterosexual couples and same-sex roommates. The input data (with integers 1–3 corresponding to categories A–C) are reproduced in Appendix B. Selected output from Inter2.sps—including cross-tabulations for the two situations, as well as the corresponding κ statistics—is shown in Appendix C, Table C2.

Conclusion

Given the difficulties of analyzing dyadic data at the individual level and the necessity to make adjustments to statistics, significance tests, and confidence intervals obtained via the SPSS graphic interface, as well as the laborious and error-prone task of restructuring the data—especially when the data set is large—the user-friendly approach adopted in *Inter1.sps* and *Inter2.sps* would prove to be useful to researchers seeking a preliminary examination of the nonindependence in their dyadic data.

AUTHOR NOTE

Correspondence concerning this article should be addressed to V. R. Alferes, Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra, Rua do Colégio Novo, Apartado 6153, 3001-802 Coimbra, Portugal (e-mail: valferes@fpce.uc.pt).

REFERENCES

- COHEN, J. (1960). A coefficient of agreement for nominal scales. *Educational & Psychological Measurement*, **20**, 37-46.
- GONZALEZ, R., & GRIFFIN, D. (1999). The correlation analysis of dyad-level data in the distinguishable case. *Personal Relationships*, **6**, 449-469.
- GRIFFIN, D., & GONZALEZ, R. (1995). Correlational analysis of dyad-level data in the exchangeable case. *Psychological Bulletin*, **118**, 430-439.
- KENNY, D. A., KASHY, D. A., & COOK, W. L. (2006). *Dyadic data analysis*. New York: Guilford.
- SHROUT, P. E., & FLEISS, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, **86**, 420-428.

NOTES

1. If the user has an individual structure data set—that is, each record refers to an individual—with two identifier variables (e.g., *dyad* for dyads and *person* for persons within dyads), he or she can easily convert this structure to a dyadic structure by running the following syntax lines:

```
delete variables person.
casestovars/id = dyad/groupby = index.
```

2. Here we describe *Inter2.sps* for estimating κ on tabulated data. If the user wants to perform the calculations directly on nontabulated raw data, he or she may run *Inter3.sps*, which is available from the authors' Web sites at <http://davidakenny.net/kkc/c2/c2.htm> or www.fpce.uc.pt/niips/inter.htm. To save space, we omit *Inter3.sps* syntax and output from the Appendixes.

SUPPLEMENTAL MATERIALS

The SPSS syntax discussed here and the data files may be downloaded as supplemental materials for this article from brm.psychonomic-journals.org/content/supplemental.

APPENDIX A

Inter1.sps

```
preserve.
set printback=on.

* MEASUREMENT OF NONINDEPENDENCE IN STANDARD DYADIC DESIGNS
  Inter1.sps - Interval measures.

* Example data: Kenny, Kashy & Cook (2006, p. 33, Table 2.1).
data list free/dyad x.1 x.2.
begin data.
1      8      6
2      5      3
3      7      2
4      8      5
5      8      7
6      5      6
7      3      4
8      8      9
9      6      7
10     2      3
end data.

*\beginning of the program.
set printback=on.

/* User Specifications /*

* Enter alpha for confidence intervals (by default, alpha=.05).
compute alpha=.05.

set printback=off.
set ovars labels onumbers values tvars labels tnumbers values.
save outfile=file1.
```

APPENDIX A (Continued)

```

/* Part 1 - Distinguishable members: Pearson correlation (r)/*

* Calculating and testing r.
descriptives variables=x.1 x.2/save.
compute zz=Zx.1*Zx.2.
compute k=1.
aggregate/outfile=*/break=k/zzsum=sum(zz)/alpha=max(alpha)/n=n.
compute r=zzsum/(n-1).
compute df=n-2.
compute t=(r*sqr(df))/(sqr(1-r**2)).
compute p=2*(1-cdf.t(abs(t),df)).

* Confidence interval for r.
compute z=(ln(1+r)-ln(1-r))/2.
compute zlow=z-idf.normal((1-alpha/2),0,1)/sqr(n-3).
compute zupp=z+idf.normal((1-alpha/2),0,1)/sqr(n-3).
compute rlow=(exp(2*zlow)-1)/(exp(2*zlow)+1).
compute rupp=(exp(2*zupp)-1)/(exp(2*zupp)+1).

* Output management.
formats all(f8.2) n df(f8.0) p(f8.3).
variable labels p"Sig. (2-tailed)" t"t Value"
  rlow"CI of r \n Lower Bound" rupp"CI of r \n Upper Bound".
oms/select tables/if commands=["variables to cases"]
  subtypes=["generated variables"]/destination viewer=no.
oms/select tables/if commands=["variables to cases"]
  subtypes=["processing statistics"]/destination viewer=no.
oms/select tables/if commands=["unianova"]
  subtypes=["between subjects factors"]/destination viewer=no.
oms/select tables/if commands=["unianova"] subtypes
  =["test of between subjects fixed effects"]/destination viewer=no.
oms/select tables/if commands=["summarize"]
  subtypes=["case processing summary"]/destination viewer=no.
summarize/tables=n r t df p rlow rupp/format=list nocasenum nototal
/title='Distinguishable Members: \n Pearson Correlation, t Test,'
+' and Confidence Interval'/cells=none.

/* Part 2 - Indistinguishable members: Intraclass correlation (ICC)/*

* Restructure file.
get file=file1.
varstocases/make dv x.1 x.2/index=person.

* Calculating and testing ICC.
unianova dv by dyad/design=dyad/outfile=effect(file2).
get file=file2.
if(source_='S2') msb=ms.
if(source_='Error') msw=ms.
if(source_='S2') df1=df.
if(source_='Error') df2=df.
compute k=1.
aggregate/outfile=*/break=k/msb=max(msb)/msw=max(msw)
  /df1=max(df1)/df2=max(df2).
compute icc=(msb-msw)/(msb + msw).
if (msb>=msw) f=msb/msw.
if (msb<msw) f=msw/msb.
if (msb>=msw) p=2*(1-cdf.f(f,df1,df2)).
if (msb<msw) p=2*(1-cdf.f(f,df2,df1)).
match files/file=*/file=file1.

* Confidence interval for ICC.
compute fl=(msb/msw)/idf.f((1-alpha/2),df1,df2).
compute fu=(msb/msw)*idf.f((1-alpha/2),df2,df1).
compute icclow=(fl-1)/(fl+1).

```

APPENDIX A (Continued)

```

compute iccupp=(fu-1)/(fu+1).

* Output management.
formats all(f8.2) df1 df2(f8.0) p(f8.3).
variable labels p"Sig. (2-tailed)" icc"ICC" f"F"
  icclow"CI of ICC \n Lower Bound" iccupp"CI of ICC \n Upper Bound"
  msb"Between Dyads \n Mean Square" msw"Within Dyads \n Mean Square".
n of cases 1.
summarize/tables=icc icclow iccupp/format=list nocasenum nototal
  /title='Indistinguishable Members: \n Intraclass Correlation'
  +' and Confidence Interval'/cells=none.
summarize/tables=msb df1 msw df2 f p/format=list nocasenum nototal
  /title='F Test for Intraclass Correlation'/cells=none.

/* Part 3 - Indistinguishable members: Pairwise correlation (rp) /*

* Restructure file.
get file=file1.
save outfile=file3.
rename variables (x.1=x.2) (x.2=x.1).
add files/file=*/file='file1'.

* Calculating and testing rp.
descriptives variables=x.1 x.2/save.
compute zz=Zx.1*Zx.2.
compute k=1.
aggregate/outfile=*/break=k/zzsum=sum(zz)/n=n.
compute rp=zzsum/(n-1).
compute sde=1/sqr(n/2).
compute z=rp/sde.
compute p=2*(1-cdf.normal(abs(z),0,1)).
match files/file=*/file=file1.

* Confidence interval for rp.
compute rpLOW=rp-(idf.normal((1-alpha/2),0,1)/sqr(n/2)).
compute rpUPP=rp+(idf.normal((1-alpha/2),0,1)/sqr(n/2)).

* Output management.
formats all(f8.2) n(f8.0) p(f8.3).
variable labels p"Sig. (2-tailed)" rp"rp" z"z Value" sde"Std. Error"
  rpLOW"CI of rp \n Lower Bound" rpUPP"CI of rp \n Upper Bound".
n of cases 1.
summarize/tables=n rp rpLOW rpUPP/format=list nocasenum nototal
  /title='Indistinguishable Members: \n Pairwise Correlation and'
  +' Confidence Interval'/cells=none.
summarize/tables=z sde p/format=list nocasenum nototal
  /title='z Test for Pairwise Correlation'/cells=none.

get file=file1/drop alpha.
save outfile=file0.sav.
set printback=on.

*/end of the program.
restore.
```

APPENDIX B
Inter2.sps

```
preserve.
set printback=on.

* MEASUREMENT OF NONINDEPENDENCE IN STANDARD DYADIC DESIGNS
  Inter2.sps - Categorical measures / Tabulated data.

* Example data: Kenny, Kashy & Cook (2006, p. 41, Table 2.4).
data list free/Member1 Member2 count.
begin data.
1      1      106
1      2       10
1      3        4
2      1       22
2      2       28
2      3       10
3      1        2
3      2       12
3      3        6
end data.

*\beginning of the program.

set printback=off.
oms/select tables/if commands=["crosstabs"]
  subtypes=["case processing summary"]/destination viewer=no.
set printback=on.

/* Distinguishable members /*
set printback=off.

* Calculating kappa.
weight off.
sort cases by member1(a) member2(a).
weight by count.
crosstabs/tables=member1 by member2/statistic=kappa.
set printback=on.

/*Indistinguishable members /*
set printback=off.

* Averaging off-diagonal cells.
save outfile=file1.
sort cases by member2(a) member1(a).
compute count2=count.
execute.
delete variables member1 member2 count.
match files/file=*/file=file1.
compute meancount=mean(count2,count).

* Calculating kappa.
weight by meancount.
crosstabs/tables=member1 by member2/statistic=kappa/count asis.
delete variables count2 meancount.
save outfile=file0.sav.
set printback=on.

*/end of the program.
restore.
```

APPENDIX C
Selected Output from Inter1.sps and Inter2.sps

Table C1
Inter1.sps: Measurement of Nonindependence in Standard Dyadic Designs (Interval Measures)

[...]

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
x.1	10	2	8	6.00	2.211
x.2	10	2	9	5.20	2.201
Valid N (listwise)	10				

[...]

Distinguishable Members: Pearson Correlation, *t* Test, and Confidence Interval

	<i>n</i>	<i>r</i>	<i>t</i> Value	<i>df</i>	Sig. (2-tailed)	CI of <i>r</i> Lower Bound	CI of <i>r</i> Upper Bound
1	10	.53	1.75	8	.119	-.16	.87

[...]

Indistinguishable Members: Intraclass Correlation and Confidence Interval

	ICC	CI of ICC Lower Bound	CI of ICC Upper Bound
1	.51	-.10	.85

[...]

F Test for Intraclass Correlation

	Between Dyads Mean Square	<i>df</i> 1	Within Dyads Mean Square	<i>df</i> 2	<i>F</i>	Sig. (2-tailed)
1	7.42	9	2.40	10	3.09	.093

[...]

Descriptive Statistics

	<i>N</i>	Minimum	Maximum	Mean	Std. Deviation
x.1	20	2	9	5.60	2.186
x.2	20	2	9	5.60	2.186
Valid N (listwise)	20				

[...]

Indistinguishable Members: Pairwise Correlation and Confidence Interval

	<i>n</i>	<i>rp</i>	CI of <i>rp</i> Lower Bound	CI of <i>rp</i> Upper Bound
1	20	.47	-.15	1.09

[...]

z Test for Pairwise Correlation

	<i>z</i> Value	Std. Error	Sig. (2-tailed)
1	1.49	.32	.136

[...]

Table C2

Inter2.sps: Measurement of Nonindependence in Standard Dyadic Designs (Categorical Measures)

[...]

/* Distinguishable members /*

Member1 * Member2 Crosstabulation

Count		Member2			Total
		1	2	3	
Member1	1	106	10	4	120
	2	22	28	10	60
	3	2	12	6	20
Total		130	50	20	200

Symmetric Measures

		Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Measure of Agreement	Kappa	.429	.054	7.720	.000
N of Valid Cases		200			

^aNot assuming the null hypothesis.

^bUsing the asymptotic standard error assuming the null hypothesis.

APPENDIX C (Continued)

[...]

/*Indistinguishable members /*

Member1 * Member2 Crosstabulation

Count

		Member2			Total
		1	2	3	
Member1	1	106.000	16.000	3.000	125.000
	2	16.000	28.000	11.000	55.000
	3	3.000	11.000	6.000	20.000
Total		125.000	55.000	20.000	200.000

Symmetric Measures

		Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Measure of Agreement	Kappa	.427	.054	7.648	.000
N of Valid Cases		200			

^aNot assuming the null hypothesis.^bUsing the asymptotic standard error assuming the null hypothesis.

[...]

(Manuscript received May 22, 2008;
revision accepted for publication July 16, 2008.)