A SAS MACRO for estimating bootstrapped confidence intervals in dyadic regression models.

Robert E. Wickham, Texas Institute for Measurement, Evaluation, and Statistics, University of Houston, TX

ABSTRACT

The actor-partner interdependence model (APIM; Kenny, Kashy, & Cook, 2006) is a popular procedure for analyzing dyadic data (e.g., married couples, twin-siblings). Recent work (Kenny & Ledermann, 2010; Wickham & Knee, 2012) demonstrates that the ratios of regression coefficients generated by the APIM are often of substantive interest to researchers. Unfortunately, the sampling distributions for these ratios are non-normal, which renders standard parametric tests of significance (e.g., *t*-test) unreliable. This talk presents a SAS MACRO for performing a non-parametric bootstrapping procedure that will provide unbiased confidence intervals for several of the key coefficient ratios estimated by the APIM. Extensions to this general model are discussed.

INTRODUCTION

Behavior researchers are often interested in phenomena which are inherently interpersonal in nature. In order to provide a more theoretically compelling and ecologically valid assessment of these processes, researchers often collect data from multiple members of intact social groups. For example, to understand the role of social support provision in close relationships, one may obtain reports of the frequency of supportive behaviors (e.g., listening to relationship partner's problems at work) and some indicator of relationship well-being (e.g., commitment to the relationship) from both spouses in a sample of married couples. Dyadic processes are often of interest in occupational or work settings. For example, it may be important to understand the nature of mentor-mentee relationships as they relate to training and subsequent job performance.

Regardless of the specific application, collecting responses from both members of a dyad introduces non-independence among scores, which violates a fundamental assumption of statistical procedures based on the general linear model (i.e., ordinary least squares regression, analysis of variance, etc.), leading to biased tests of significance for model parameters. As a result, an application of mixed-effects modeling known as the actor-partner interdependence model (APIM; Kenny, 1996) was developed, which accounts for non-independence among dyad member responses and provides unbiased significance tests. In the APIM, *actor* regression coefficients (slopes) describe the association between one member's predictor (e.g., man's social support provision) and his/her own outcome (e.g., man's relationship commitment), whereas *partner* coefficients describe the association between the predictor for one's dyadic partner (e.g., woman's social support provision) and his/her own outcome (e.g., man's relationship commitment). Finally, the *actor*×*partner* interaction describes how the unique combination of each dyad member's standing on the predictor relates to the outcome.

Recent work has shown that the ratio of *actor*:*partner* coefficients (i.e., *k*), *actor:actor*×*partner* coefficients (i.e., *c*), and *partner:actor*×*partner* coefficients (i.e., *h*) correspond to components of a major theory describing dyadic interaction, known as interdependence theory (Wickham & Knee, 2012). Unfortunately, the sampling distribution of these ratios is non-normal, and researchers need a method for obtaining confidence intervals describing the range of likely values for each ratio. The ***%kchboot*** MACRO utilizes a resampling procedure to derive bootstrapped estimates for these ratios.

MACRO PROCEDURE

The *%**kchboot*** MACRO must manipulate the format of the input data in order to perform the resampling procedure, after which the data must be re-arranged back into univariate format for analysis. The bootstrapped estimates are then compiled and confidence intervals are computed and reported. Each of these steps and the associated code are described below.

UNIVARIATE TO MULTIVARIATE DYAD FORMAT

The mixed effects modeling procedure used to estimate actor, partner and *actor*×*partner* regression coefficients (i.e., PROC MIXED), requires data in univariate format with a unique ID variable for each dyad (*idvar*), as well as a variable used to distinguish between members in the same dyad (*distvar*; e.g., gender; mentor-mentee status). After sorting the input dataset by *idvar* and *distvar*, new numeric IDs are generated for each dyad, the data are re-arranged into multivariate format with each row containing predictor and outcome scores for each member of the dyad. This is accomplished by creating temporary datasets for each dyad member type, however the TRANSPOSE procedure could also have been used.

```
[%MACRO kchboot (data, dyads, nboot, idvar, distvar, memb_a, memb_b);
*Sort Original Dataset;
proc sort data = &data;
by &idvar &distvar;
run;
*Create new IDs;
data newid;
set &data;
if mod(_n_,2) ne 0 then nid = (_n_+ 1) / 2;
else if mod(_n_,2) eq 0 then nid = _n_ / 2;
run;
*Convert original dataset to mv format;
data long_a;
set newid;
where &distvar eq "&memb_a";
rename actor = pred_a outcome = outcome_a;
keep nid actor outcome;
run;
data long_b;
set newid;
where &distvar eq "&memb_b";
rename actor = pred_b outcome = outcome_b;
keep nid actor outcome;
run;
data orig_mv;
merge long_a long_b;
by nid;
run;
```

RESAMPING PROCEDURE

Next, *k* bootstrap samples of size *n* are drawn (with replacement) from the sample dataset, where *k* is specified by the user (*nboot*) and *n* = the number of dyads present in the original sample dataset. The resampled datasets are stored in a single dataset, with a resample ID, for subsequent analysis.

```
*Create resampled dataset;
%DO h = 1 %TO &nboot;
data bootsamp_mv;
      sampid = &h;
      do i = 1 to &dyads;
            x = int(ranuni(-1) * &dyads)+1;
            set orig_mv
                  nobs = nobs
                  point = x;
            output;
      end;
      stop;
run;
proc append base = alldata data = bootsamp_mv;run;
%END;
```

MULTIVARIATE TO UNIVARIATE DYAD FORMAT

In order to conduct APIM analysis on the resampled datasets, the data must be transformed back into univariate format. Dummy variables designating each dyad member's role must also be created.

```
*Convert resampled dataset to univariate format;
data a_memb;
set alldata;
dum_a = 1; dum_b = 0; eff_a = 1; distvar = "memb_a";
rename outcome_a = outcome pred_a = actor pred_b = partner;
keep sampid nid i distvar dum_a dum_b eff_a outcome_a pred_a pred_b;
run;
data b_memb;
set alldata;
dum_a = 0; dum_b = 1; eff_a = -1; distvar = "memb_b";
rename outcome_b = outcome pred_b = actor pred_a = partner;
keep sampid nid i distvar dum_a dum_b eff_a outcome_b pred_a pred_b;
run;
data bootsamp_uni;
set a_memb b_memb;
run;
proc sort data = bootsamp_uni;
by sampid i distvar;
run;
```

APIM ANALYSIS

Next, APIM analysis is conducted on each resampled dataset, and the *actor*, *partner*, and *actor*×*partner* regression coefficients are stored in a SAS dataset for secondary analysis/processing. Two independent analyses were performed for each sample. In the pooled analyses, the magnitude of *actor*, *partner*, and *actor*×*partner* coefficients are assumed to be equal for each dyad member type (i.e., men and women), in the separate analysis *actor*, *partner*, and *actor*×*partner* coefficients are estimated independently for each dyad member type.

```
*Pooled Parameters;
ods select none;
proc mixed data = bootsamp_uni noclprint noitprint;
by sampid; class distvar;
model outcome = eff_a|actor|partner/notest s;
repeated distvar / sub = i type = un;
ods output solutionf = pooled_mv (keep = sampid Effect Estimate);
run;
proc transpose data = pooled_mv out = pooled_uni (keep = col3 col5 col7);
by sampid; var estimate; run;
data pooled_param; set pooled_uni;
k = col3/col5; c = col5/col7; h = col3/col7;keep k c h;
run;
*Memb_a Parameters;
proc mixed data = bootsamp_uni noclprint noitprint;
by sampid; class distvar;
model outcome = dum_b|actor|partner/notest s;
repeated distvar / sub = i type = un;
ods output solutionf = a_mv (keep = sampid Effect Estimate);
run;
proc transpose data = a_mv out = a_uni (keep = col3 col5 col7);
by sampid; var estimate; run;
data a_param; set a_uni;
k = col3/col5; c = col5/col7; h = col3/col7;keep k c h;
run;
*Memb_b Parameters;
proc mixed data = bootsamp_uni noclprint noitprint;
by sampid; class distvar;
model outcome = dum_a|actor|partner/notest s;
repeated distvar / sub = i type = un;
ods output solutionf = b_mv (keep = sampid Effect Estimate);
run;
proc transpose data = b_mv out = b_uni (keep = col3 col5 col7);
by sampid; var estimate; run;
data b_param;set b_uni;
k = col3/col5; c = col5/col7; h = col3/col7;keep k c h;
run;
```

OUTPUTTING ESTIMATES

Finally, 80, 90 and 95% confidence intervals are printed for pooled and separate estimates.

```
*Output Ratios;
ods select all;
title "Pooled Ratio Parameters";
proc univariate data = pooled_param;
var k c h;
histogram k c h;
output out = ci95_pooled pctlpts = 2.5 50 97.5 pctlpre = k c h pctlname =
_lower_95CI _median_ _upper_95CI;run;
proc print data = ci95_pooled;run;
title;
title "Member 'A' Ratio Parameters";
proc univariate data = a_param;
var k c h;
histogram k c h;
output out = ci95_a pctlpts = 2.5 50 97.5 pctlpre = k c h pctlname =
_lower_95CI _median_ _upper_95CI;run;
proc print data = ci95_a;run;
title;
title "Member 'B' Ratio Parameters";
proc univariate data = b_param;
var k c h;
histogram k c h;
output out = ci95_b pctlpts = 2.5 50 97.5 pctlpre = k c h pctlname =
_lower_95CI _median_ _upper_95CI;run;
proc print data = ci95_b;run;
title;
%MEND;
```

WORKED EXAMPLE

The following section demonstrates the application of ***%kchboot***. Both members of 71 heterosexual dyads provided reports of general life stress (M = 1.85, SD = 0.69) and life satisfaction (M = 5.06, SD = 1.24) using Likert-type response scales. Pooled APIM analyses revealed a significant actor effect, suggesting that higher levels of stress were associated with lower life satisfaction. None of the remaining main effects or interactions reached significant. However, separate analyses for each gender revealed an unexpected partner effect for men, suggesting that men with partners who report higher levels of stress, tend to report higher levels of relationship satisfaction. Pooled results are provided in Table 1.

Table 1.

Parameter	***Estimate***	***S.E.***	***t-value***
Intercept	5.04	0.10	51.09**
Gender	-0.17	0.09	-1.81†
Actor	-0.85	0.15	-5.88**
Partner	0.18	0.15	1.21
AxG	-0.10	0.15	-0.67
PxG	0.20	0.15	1.35
AxP	0.33	0.24	1.38
AxPxG	0.14	0.22	0.62
σ^2_M	1.03	0.18	5.79**
σ^2_W	1.46	0.25	5.74**
$\sigma_{M,W}$	0.09	0.15	0.54

To obtain point estimates and bootstrapped estimates of the *actor*, *partner*, and *actor×partner* regression coefficients, the data were submitted to the *%kchboot* MACRO. However, the variables must first be renamed so that the actor predictor (stress) is named 'actor', the partner predictor is named 'partner', and the outcome (life satisfaction) is named 'outcome'. The MACRO is called using the following command:

```
%kchboot(data = mydata, dyads = 71, nboot = 1000, idvar = CID, distvar =
gender, memb_a = boy, memb_b = girl);
```

The final step of the MACRO executes PROC UNIVARIATE to produce confidence intervals for each ratio. The actual output is too extensive to list here, however a summary table is provided below:

Parameter	Gender	Lower 2.5%	Lower 5%	Lower 10%	Point Est.	Upper 10%	Upper 5%	Upper 2.5%
k (AC:PC)	Men	-25.68	-12.80	-6.16	-0.30	5.03	11.35	21.73
	Women	-0.93	-0.35	-0.10	0.52	1.75	2.78	5.38
h (AC:JC)	Men	-7.34	-3.78	-1.87	-0.08	2.02	3.63	5.33
	Women	-21.12	-7.88	-3.73	-0.86	2.39	5.11	12.43
c (PC:JC)	Men	-18.26	-6.42	-3.33	-0.24	2.50	5.72	11.32
	Women	-12.23	-5.11	-2.81	-0.37	1.86	4.14	8.70

CONCLUSIONS

The ***%kchboot*** MACRO provides non-parametric confidence intervals for ratios of regression coefficients in the actor-partner interdependence model. These ratios, and the associated confidence intervals, convey important information regarding the nature of dyadic interdependence present for a set of constructs. The bootstrap resampling component of the MACRO may be adapted to support triads (i.e., groups of 3), or larger groups with distinguishable roles, or other dyadic analysis models (e.g., common-fate model).

REFERENCES:

Kenny, D. A. (1996). Models of non-independence in dyadic research. *Journal of Social and Personal Relationships, 13,* 279-294.

Kenny, D. A., & Ledermann, T. L. (2010). Detecting, measuring, and testing dyadic patterns in the actor-partner interdependence model. *Journal of Family Psychology, 24,* 359-366. doi: 10.1037/a0019651

Wickham, R. E., & Knee, C. R. (2012). Interdependence theory and the actor-partner interdependence model: Where theory and method converge. *Personality and Social Psychology Review, 16,* 375-393.

ACKNOWLEDGMENTS

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Robert E. Wickham, Ph.D.
Texas Institute for Measurement, Evaluation, and Statistics
Department of Psychology
University of Houston
126 Heyne Building
Houston, TX 77204-5502
713-743-8500
robert.wickham@times.uh.edu

* *