GEN_ETA2: A SAS® Macro for Computing the Generalized Eta-Squared Effect Size Associated with Analysis of Variance Models

Patricia Rodríguez de Gil, Thanh Pham, Patrice Rasmussen, Anh Kellermann, Jeanine Romano, Yi-Hsin Chen, Jeffrey D. Kromrey
University of South Florida

INTRODUCTION

- Measures of effect size are recommended for communicating information on the strength of relationships between variables
- Effect size information supplements the reject/fail to reject decision obtained in statistical hypothesis testing
- The choice of an effect size for ANOVA models can be confusing because indices may differ depending on the research design as well as the magnitude of the effect

PURPOSE

- This paper provides a SAS® macro for computing the generalized eta squared effect size (Olejnick & Algina, 2003), associated with analysis of variance models by utilizing data from PROC GLM ODS tables
- The paper provides the macro programming language, as well as results from an executed example of the macro

EFFECT SIZES COMMONLY USED WITH WITH ANALYSIS OF VARIANCE MODELS

ETA SQUARED (η^2)

$$\eta^2 = \frac{SS_{effect}}{SS_{Total}}$$

PARTIAL ETA SQUARED η_p^2

$$\eta_p^2 = \frac{SS_{effect}}{SS_{effect} + SSs/cells}$$

- Both η^2 and η_p^2 sample effect size estimates represent the proportion of variability in the dependent variable that is associated with variability in an independent variable.
- However, these statistics are positively biased as point estimates of the population effect size.

OMEGA SQUARED (ω^2)

$$\omega^2 = \frac{SS_{effect} - (k-1)MSerror}{SS_{Total} + MSerror}$$

PARTIAL OMEGA SQUARED (ω_p^2)

$$\omega_p^2 = \frac{SS_{effect} - (k-1)MSerror}{SS_{Total} + (N-k-1)MSerror}$$

- To obtain a relatively unbiased estimate of the variance explained in the population by an independent variable, ω^2 can be calculated
- Unlike eta-squared, ω^2 takes random error (MS_{error}) into account.
- Thus ω^2 values will be smaller than η^2 , with more noticeable differences occurring with smaller samples and/or research designs that include more independent variables

GENERALIZED ETA SQUARED (η_G^2)

- Proposed by Olejnick and Algina, 2003
- Its value is not influenced by the study design; consequently, it is comparable across designs.
- The Generalized Eta-Squared statistic is estimated:

$$\eta_G^2 = \frac{SS_{effect}}{\delta(SS_{effect}) + \sum_{Meas} SS_{Meas} + \sum_K SS_K}$$

- where SS_{effect} is the sum of squares for the effect of interest,
- $\delta = 1$ if the effect is a manipulated factor (and is zero otherwise),
- the SS_{Meas} are the sums of squares for all sources of variance that involve measured factors (rather than manipulated factors) but do not include subjects, and
- the SS_K are the sums of squares for all sources of variance that involve subjects

SOFTWARE LIMITATIONS

- PROC ANOVA and PROC GLM do not provide the generalized eta-squared effect size (with the exception of single factor models, for which η^2 is equivalent to η_G^2
- For SAS® users to follow the APA Task Force recommendations, for reporting effect sizes, extra work is required.
- ModelANOVA and OverallANOVA ODS tables can be used to output the sum-of-squares needed to calculate generalized eta-squared either by hand or through a data step
- However, computation formulas for estimating η_G^2 can be confusing

MODELS FOR WHICH η_G^2 IS COMPUTED

_	_	Within-Subjects Factors		
-		0	1	2
Between-Subjects Factors	0	N-A	✓	✓
	1	✓	✓	✓
	2	✓	✓	✓
	> 2	✓	✓	✓

- The macro computes η_G^2 for ANOVA designs indicated above
- It is intended for designs with at least one categorical independent variable
- Distinction of the study factors guide its estimation
  - Measured factors
  - Manipulated factors

GEN_ETA2 MACRO PARAMETERS

```
DATA       = _LAST_      User provided
CLASS      =
MODEL      = NONE
REPEATED   = NONE        Defaults
MEASURED   = NONE
DEPENDENT  = Y
```

GEN_ETA2 MACRO SAMPLE DATA

BETWEEN-SUBJECTS ONLY

```
DATA ONE;
 INPUT SCORE GROUP;
 SUBJECT = _N_;
CARDS;
 17 1
 23 1
 20 1
 14 1
 12 1
 18 2
 22 2
 11 2
 21 2
 19 2
;
RUN;
```

BETWEEN AND WITHIN-SUBJECTS DESIGN

```
DATA ONE;
 INPUT X1 X2 X3 GROUP SEX INTENSITY;
 SUBJECT = _N_;
CARDS;
 1 2 3 1 1 1
 1 1 1 1 1 2
 2 2 5 1 2 2
 3 3 3 1 2 1
 3 2 1 1 2 1
 6 5 4 2 1 2
 4 5 4 2 2 2
 3 6 7 2 1 1
 3 4 8 2 2 1
;
RUN;
```

TWO WITHIN-SUBJECTS DESIGN

```
DATA ONE
 INPUT STUDENT COOP1 COOP2 AVOID1 AVOID2
PEER1 PEER2;
CARDS;
1 31 33 21 27 28 30
2 31 25 15 25 30 32
3 16 35 26 33 17 32
4 27 26 22 31 21 31
5 32 26 29 32 29 35
6 32 33 16 33 20 27
7 26 32 14 27 19 19
8 19 26 14 29 10 29
;
RUN;
```

GEN_ETA2 MACRO EXECUTION

Sample macro calls:

%***gen_eta2*** (data = one, class= group , model = group, dependent=score);

%***gen_eta2*** (data = one, repeated = %STR(time 2, domain 3), dependent=coop1 avoid1 peer1 coop2 avoid2 peer2);

OUTPUT EXAMPLES

```
Generalized Eta-Squared Values

Source           SS          Hypothesis Type        generalized Eta2

WithinSubjects

  Domain       214.0833     3                        0.2045
```

Output 1. Output Example for One Within-Subjects Factor Design

```
Generalized Eta-Squared Values

Source            SS        Hypothesis Type   Generalized Eta2

WithinSubjects

    time           3              0.0818
    time*domain   109.2917       3               0.0849
```

Output 2. Output E xample for Two Within-Subjects Factor Design

```
Generalized Eta-Squared Values

Source                            SS  Hypothesis Type  Generalized Eta2
BetweenSubjects
  sex                         0.6806          3                 0.0140
  group                      58.6806          3                 0.5466
  intensity                   0.0139          3                 0.0003
  group*sex                   1.1250          3                 0.0231
  sex*intensity               0.0139           3                 0.0003
  group*intensity             1.1250          3                 0.0226
  group*sex*intensity         0.1250          3                 0.0026

WithinSubject
  time                           15.1944       3                 0.2379
  time*sex                        4.1944       3                 0.0862
  time*group                      1.6944       3                 0.0336
  time*intensity                  2.0278       3                 0.0400
  time*group*sex                  2.5833       3                 0.0531
  time*sex*intensity              9.3611       3                 0.1924
  time*group*intensity            8.0833       3                 0.1424
  time*group*sex*intensity        5.0833       3                 0.1045
```

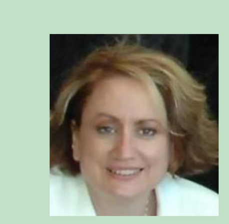
Output 3. Output example for Three Between Subjects, Two-Within Subjects Design

RESEARCH TEAM

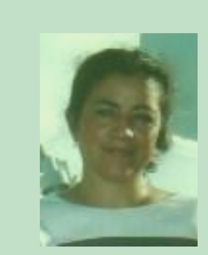
CONTACT US!

Dr. Jeffrey Kromrey
kromrey@usf.edu

Dr. Yi-Hsin Chen
ychen5@usf.edu



Dr. Jeanine Romano
romano@usf.edu



Patricia Rodriguez
Doctoral Candidate
prodrig6@usf.edu

Thanh Pham
Doctoral Candidate
tvpham2@usf.edu

Anh Kellermann
Doctoral Student
napham@mail.usf.edu

Patrice Rasmussen
Doctoral Student
prasmussen@usf.edu