What is the process for creating an interaction term in SUDAAN?

Authored by

stats writer

July 1, 2024

The process for creating an interaction term in SUDAAN involves first identifying the two or more variables that will be included in the interaction term. These variables should be categorical or continuous and have a meaningful relationship. Then, the user must specify the type of interaction (e.g. product, sum, difference) and the order of the interaction (e.g. first or second order). The interaction term is then calculated by multiplying or combining the two variables according to the specified type and order. Finally, the interaction term can be included in statistical models and analyses to assess its impact on the outcome variable.

How do I create an interaction term in SUDAAN? | SUDAAN FAQ

In proc regress, proc rlogist and proc survival, you can use
a * between two variables (such as two categorical variables or one categorical and one continuous variable) to create an interaction term on the model statement. However, you
cannot do this with two continuous variables; you need to create the interaction term in a data
step before running the model. For example, srsex and racehpra
are categorical variables. In the example below, we create the
interaction term between srsex and racehpra.

```
proc regress data=temp1 filetype=sas design = jackknife;
```

```
weight rakedw0;
jackwgts rakedw1--rakedw80 / adjjack=1;
model ab23 = srsex racehpra srsex*racehpra;
subgroup srsex racehpra;
levels 2 4;
run;
```

```
Number of observations read : 55428 Weighted count: 23847415
Observations used in the analysis : 1000 Weighted count: 466228
Denominator degrees of freedom : 80

Maximum number of estimable parameters for the model is 8
Weighted mean response is 42.854796

Multiple R-Square for the dependent variable AB23: 0.081618
Variance Estimation Method: Replicate Weight Jackknife
Working Correlations: Independent
Link Function: Identity
Response variable AB23: AB23
```

Independent P-value

Variables and Beta T-Test

Effects Coeff. SE Beta T-Test B=0 B=0

Intercept 47.87 1.81 26.51 0.0000

SRSEX

MALE 1.63 2.41 0.68 0.5011

FEMALE 0.00 0.00 . .

RACEHPRA

LATINO -9.67 2.04 -4.75 0.0000

PACIFIC ISLANDER 4.32 6.03 0.72 0.4757

AIAN -3.47 2.68 -1.29 0.1992

ASIAN 0.00 0.00 . .

SRSEX, RACEHPRA

MALE, LATINO 3.51 3.09 1.14 0.2596

MALE, PACIFIC

ISLANDER -5.64 7.27 -0.78 0.4398

MALE, AIAN -0.39 3.49 -0.11 0.9108

MALE, ASIAN 0.00 0.00 . .

FEMALE, LATINO 0.00 0.00 . .

FEMALE, PACIFIC

ISLANDER 0.00 0.00 . .

FEMALE, AIAN 0.00 0.00 . .

FEMALE, ASIAN 0.00 0.00 . .

Contrast Degrees

of P-value

Freedom Wald F Wald F

OVERALL MODEL 8 960.50 0.0000

MODEL MINUS

INTERCEPT 7 8.59 0.0000

INTERCEPT . . .

SRSEX . . .

RACEHPRA . . .

SRSEX * RACEHPRA 3 1.12 0.3457

In the next example, we will create an interaction term using a categorical

variable, racehpra, and a continuous variable, ae13.

```
proc regress data=temp1 filetype=sas design = jackknife;
weight rakedw0;
```

jackwgts rakedw1--rakedw80 / adjjack=1;
model ab23 = ae13 racehpra ae13*racehpra;
subgroup racehpra;
levels 4;
run;

Number of observations read : 55428 Weighted count: 23847415
Observations used in the analysis : 322 Weighted count: 158239
Denominator degrees of freedom : 80

Maximum number of estimable parameters for the model is 8
Weighted mean response is 40.890349

Multiple R-Square for the dependent variable AB23: 0.058683
Variance Estimation Method: Replicate Weight Jackknife
Working Correlations: Independent
Link Function: Identity
Response variable AB23: AB23

Independent P-value

Variables and Beta T-Test

Effects Coeff. SE Beta T-Test B=0 B=0

Intercept 46.99 3.16 14.87 0.0000

AE13 -0.15 1.13 -0.14 0.8917

RACEHPRA

LATINO -5.02 3.64 -1.38 0.1715

PACIFIC ISLANDER -9.37 7.35 -1.28 0.2057

AIAN -4.19 4.37 -0.96 0.3396

ASIAN 0.00 0.00 . .

AE13, RACEHPRA

1, LATINO -0.89 1.36 -0.66 0.5113

1, PACIFIC

ISLANDER 4.09 1.65 2.49 0.0150

1, AIAN -0.47 1.45 -0.33 0.7435

1, ASIAN 0.00 0.00 . .

Contrast Degrees

of P-value

Freedom Wald F Wald F

OVERALL MODEL 8 257.15 0.0000

MODEL MINUS

INTERCEPT 7 4.35 0.0004

INTERCEPT . . .

AE13 . . .

RACEHPRA 3 0.93 0.4289

AE13 * RACEHPRA 3 4.76 0.0042

In the example below, we try to use two continuous variables, ae13 and ae14, to create the interaction term. The error message

that was displayed in the log is shown below. If you want to include this

type of interaction term in your model, you will need to create it in a data

step before running the procedure and include it on the model statement.

```
proc regress data=temp1 filetype=sas design = jackknife;
weight rakedw0;
jackwgts rakedw1--rakedw80 / adjjack=1;
model ab23 = ae13 ae14 ae13*ae14;
```

run;

SEMANTIC ERROR

:

(Message 452)

At most one continuous variable is allowed

in each term of the RHS of a MODEL statement