

Electronic Supplement: Logistic GEE syntax for SPSS, SAS, and Stata

Huh, D., Flaherty, B. P., & Simoni, J. M. (2012). Optimizing the analysis of adherence interventions using logistic generalized estimating equations. *AIDS and Behavior*, 16, 422–431. doi:10.1007/s10461-011-9955-5

SPSS syntax

```
/* Format data for longitudinal analysis */

GET FILE="rctdata.sav"
/KEEP id tx1 tx2 y0 y1 y2 y3.

** Reformat outcome data into "long" format so that each participant
** has multiple rows of data, one for each assessment. Data must be
** in long (aka. univariate) format for GEE analysis.
VARSTOCASES
/MAKE y from y0 y1 y2 y3
/INDEX = time.

** Adjust time variable so baseline=0.
recode time (1=0) (2=1) (3=2) (4=3).

** The logistic GEE procedure in SPSS cannot accommodate decimals, so
** the outcome must be converted to a fraction with a numerator (p)
** and a denominator (n).
COMPUTE n = 1000000.
COMPUTE p = rnd(y * n).

/* Planned contrasts */

** Logistic GEE analysis w/ planned contrasts.
GENLIN p OF n BY tx1 tx2 time (ORDER=DESCENDING)
/MODEL tx1 tx2 time time*tx1 time*tx2
DISTRIBUTION=binomial LINK=logit
/REPEATED SUBJECT=id WITHINSUBJECT=time
CORRTYPE=ar(1) ADJUSTCORR=no COVB=robust
/PRINT cps modelinfo solution (exponentiated) summary.

/* Piecewise growth curves */

** Create piecewise growth curve.
RECODE time (0=0) (1=1) (2=2) (3=3) INTO pre.
RECODE time (0=0) (1=0) (2=1) (3=2) INTO post.

** Logistic GEE analysis w/ piecewise growth curves.
GENLIN p OF n BY tx1 tx2 (ORDER=DESCENDING) WITH pre post
/MODEL tx1 tx2 pre post
tx1*pre tx1*post tx2*pre tx2*post
DISTRIBUTION=binomial LINK=logit
/REPEATED SUBJECT=id WITHINSUBJECT=pre*post
CORRTYPE=ar(1) ADJUSTCORR=no COVB=robust
/PRINT cps modelinfo solution (exponentiated) summary.
```

SAS syntax

```
/* Format data for longitudinal analysis */

DATA rctdata;
  INFILE 'rctdata.csv' FIRSTOBS=2 DLM=', ' DSD;
  INPUT id tx1 tx2 y0 y1 y2 y3;
RUN;

** Reformat outcome data into "long" format so that each participant;
** has multiple rows of data, one for each assessment. Data must be;
** in long (aka. univariate) format for GEE analysis.;
PROC TRANSPOSE DATA=rctdata OUT=longrctdata;
  BY id;
  VAR y0 y1 y2 y3;
RUN;

DATA assignmentdata;
  SET rctdata;
  KEEP id tx1 tx2;
RUN;

** The logistic GEE procedure in SAS cannot accommodate decimals, so;
** the outcome must be converted to a fraction with a numerator (p);
** and a denominator (n);
DATA outcomedata;
  SET longrctdata (rename=(col1=y));
  time=input(substr(_name_, 2), 5.);
  n = 1000000;
  p = round(y * n, 1);
  KEEP id time n p;
RUN;

DATA mergedata;
  MERGE outcomedata assignmentdata;
  BY id;
RUN;

/* Planned contrasts */

** Create categorical contrast terms (Baseline = Reference);
DATA contrastdata;
  SET mergedata;
  time1=0;
  time2=0;
  time3=0;
  IF (time=1) THEN time1=1;
  IF (time=2) THEN time2=1;
  IF (time=3) THEN time3=1;
RUN;

** Logistic GEE analysis;
PROC GENMOD DATA=contrastdata;
  CLASS id;
```

```

MODEL p/n = tx1 tx2 time1 time2 time3
           tx1*time1 tx1*time2 tx1*time3
           tx2*time1 tx2*time2 tx2*time3 /DIST=binomial;
REPEATED SUBJECT=id / TYPE=ar CORRW;
ESTIMATE 'tx1*time1' tx1*time1 1 / EXP;
ESTIMATE 'tx1*time2' tx1*time2 1 / EXP;
ESTIMATE 'tx1*time3' tx1*time3 1 / EXP;
ESTIMATE 'tx2*time1' tx2*time1 1 / EXP;
ESTIMATE 'tx2*time2' tx2*time2 1 / EXP;
ESTIMATE 'tx2*time3' tx2*time3 1 / EXP;
CONTRAST 'tx1*time (Omnibus)' tx1*time1 1, tx1*time2 1, tx1*time3 1 / WALD;
CONTRAST 'tx2*time (Omnibus)' tx2*time1 1, tx2*time2 1, tx2*time3 1 / WALD;
RUN;

```

```
/* Piecewise growth curves */
```

```
** Create piecewise growth curve terms;
```

```
DATA growthcurvedata;
  SET mergedata;
  pre = time;
  IF (pre=0) THEN post = 0;
  IF (pre=1) THEN post = 0;
  IF (pre=2) THEN post = 1;
  IF (pre=3) THEN post = 2;
RUN;
```

```
** Logistic GEE analysis;
```

```
PROC GENMOD DATA = growthcurvedata;
  CLASS id;
  MODEL p/n = tx1 tx2 pre post
           tx1*pre tx1*post tx2*pre tx2*post /DIST=binomial;
REPEATED SUBJECT=id / TYPE=ar CORRW;
ESTIMATE 'tx1*pre' tx1*pre 1 / EXP;
ESTIMATE 'tx1*post' tx1*post 1 / EXP;
ESTIMATE 'tx2*pre' tx2*pre 1 / EXP;
ESTIMATE 'tx2*post' tx2*post 1 / EXP;
CONTRAST 'tx1*time (Omnibus)' tx1*pre 1, tx1*post 1 / WALD;
CONTRAST 'tx2*time (Omnibus)' tx2*pre 1, tx2*post 1 / WALD;
RUN;
```

Stata syntax

```
/* Format data for longitudinal analysis */
```

```
use rctdata.dta, clear
keep id tx1 tx2 y0 y1 y2 y3
```

```
** Reformat outcome data into "long" format so that each participant
** has multiple rows of data, one for each assessment. Data must be
** in long (aka. univariate) format for GEE analysis
reshape long y, i(id) j(time)
```

```

/* Planned contrasts */

** Create contrasts of time (Baseline = Reference)
generate time1 = (time==1)
generate time2 = (time==2)
generate time3 = (time==3)

** Create treatment by time interactions
generate tx1Xtime1 = tx1 * time1
generate tx1Xtime2 = tx1 * time2
generate tx1Xtime3 = tx1 * time3

generate tx2Xtime1 = tx2 * time1
generate tx2Xtime2 = tx2 * time2
generate tx2Xtime3 = tx2 * time3

** Logistic GEE analysis
xtgee y tx1 tx2 time1 time2 time3 ///
      tx1Xtime1 tx1Xtime2 tx1Xtime3 ///
      tx2Xtime1 tx2Xtime2 tx2Xtime3, ///
      i(id) family(binomial) link(logit) corr(ar1) t(time) robust eform

** Omnibus tests of intervention effects
testparm tx1Xtime1 tx1Xtime2 tx1Xtime3
testparm tx2Xtime1 tx2Xtime2 tx2Xtime3

/* Piecewise growth curves */

** Create piecewise growth curves
recode time (0=0) (1=1) (2=2) (3=3), generate(pre)
recode time (0=0) (1=0) (2=1) (3=2), generate(post)

** Create treatment by time interactions
generate tx1Xpre = tx1 * pre
generate tx1Xpost = tx1 * post

generate tx2Xpre = tx2 * pre
generate tx2Xpost = tx2 * post

** Logistic GEE analysis
xtgee y tx1 tx2 pre post ///
      tx1Xpre tx2Xpre tx1Xpost tx2Xpost, ///
      i(id) family(binomial) link(logit) corr(ar1) t(time) robust eform

** Omnibus tests of intervention effects
testparm tx1Xpre tx1Xpost
testparm tx2Xpre tx2Xpost

```