Paper 216-2011

Contrast Statements for Binary Responses: Comparing and Contrasting with PROC GLIMMIX

Jeff Gossett, UAMS Pediatrics, Little Rock, AR, USA

Mary Aitken, UAMS Pediatrics, Little Rock, AR, USA

Beverly K. Miller, UAMS Pediatrics, Little Rock, AR, USA

Christopher Swearingen, UAMS Pediatrics, Little Rock, AR, USA

ABSTRACT

SAS® provides many useful options for post-hoc estimation within the GLIMMIX procedure. When the response is binary, PROC GLIMMIX can produce estimates that are link transformed, proportions, and odds ratios, depending on the options selected. Using data from a community intervention study, this paper demonstrates how to use post-hoc estimation statements in PROC GLIMMIX to produce estimates and test hypotheses to evaluate the intervention.

INTRODUCTION

As our data source, we use the StrikeOut Community Intervention Study. The main outcome of interest for this study is whether 4 to 7 year old subjects are properly restrained in cars based on American Academy of Pediatrics (AAP) standards. The primary hypothesis is that the community intervention improves adherence to the AAP standards for proper restraint of children. The study collected data in 10 control and 10 intervention communities in which a tailored educational program was conducted; 3 communities were assessed as a control and a year later as intervention. In each community, there were pre and post intervention observations of restraint use. Interventions were conducted in 4 states – Alabama, Arkansas, Illinois, and Indiana, which have different laws regarding restraint use. Because entire communities were randomized to intervention status (as opposed to individuals within communities being randomized) and the interventions were applied at the community level, a technique that adjusts for clustering is needed. PROC GLIMMIX allows the use of both fixed and random effects. Our model addresses the clustering by including a random intercept term that varies by community. Murray and Van Horn describe various models, some of which contain multiple random effects. However, this paper does not address contrasts involving random effects.

METHODOLOGY

The objective of the analysis was to use a statistical model to assess the success of the intervention. The outcome is a dichotomous variable that indicates whether the child was properly secured in either a car seat or booster seat as appropriate according to AAP standards. For various combinations of the design variables, we would like to predict the probability that a child will be properly restrained and to make statistical inferences. Design variables included: state (AL - Alabama, AR-Arkansas, IN-Indiana, IL-Illinois), treatment group (intervention or control), time period (pre or post), as well as all the 2- and 3-way interactions between these variables. The analysis is complicated by subjects being randomized to treatment group (intervention or control) at the community rather than individual level. A random community effect was included in the model to account for possible correlation between subjects from the same community. In terms of statistical language, a generalized linear mixed model (GLMM) with a binomial link was used to model to the log odds of the probability that a child was restrained according to AAP guidelines. The GLMM allows us to include a random community effect. Two assessments were made within each community (pre/post). Fixed effects included the main effects: treatment group (test: 0=control, 1= intervention), time (post: 0=pre, 1=post), state (state: 1=AL, 2=AR, 3=IL, 4=IN), All 2- and 3-way interactions between treatment group, time, and state were included. The outcome AAP was coded 1=meets AAP standards, 0=does not.

After sorting the data by community, the basic SAS code for this model was:

```
ods graphics on;
proc glimmix data=pass2 oddsratio plots=oddsratio;
    class POST TEST state community;
    model AAP=POST|TEST|state / solution dist=binary oddsratio;
    random intercept /sub=community;
run;
```

The `oddsratio` option on the PROC GLIMMIX statement, requests odds ratio calculations for main effects. The `AAP=POST|TEST|STATE` on the model statement is expanded by SAS to include all main effects (`POST`, `TEST`, and `STATE`), two-way interactions (`POST*TEST`, `POST*STATE`, and `TEST*STATE`), and the three-way interaction (`POST*TEST*STATE.`)

OBJECTIVES OF PAPER

PROC GLIMMIX has capabilities to answer questions such as the following. The following issues were addressed in our analysis:

1. For each combination of state, test, and time we might want to estimate the probability that a child will meet AAP standards. (demonstrates `ilink` option on `lsmeans` statement)
2. Within each state and treatment group, did the probability of meeting AAP change from pre- to post-assessment? (demonstrates hypothesis testing using slice option of `lsmeans` statement)
3. By state, compute the odds of meeting AAP after the treatment versus before. At post assessment, compute the odds of meeting AAP in the intervention group versus the control group by state. (demonstrates odds ratio calculations using `lsmeans` statement with `diffs`, `or`, and `cl` options.)
4. Overall calculations – combining state results. (demonstrates `estimate` with or and `cl` options.)

One by one, the analysis objectives will be addressed. Some objectives can be met easily using `lsmeans` statements. Others require more complicated `estimate` statements.

Objective 1: estimate probability of adherence to AAP standards by Post*Test*State.

This objective can be addressed using the LSMEANS statement as follows:

```
lsmeans post*test*state post*test / ilink cl;
```

By default `lsmeans` produces estimates on the logit scale. The ILINK option on the LSMEANS statement requests that the estimates be transformed back to the scale of the original data (i.e. proportion meeting AAP). The LSMEANS output will include estimates of the probability of meeting AAP by each combination of time, test, and state and time and test   The CL option requests confidence intervals for the estimates. The overall results (LSMEANS post*test), averaging over states, are omitted for brevity.

Objective 1 output – Link transformed and Data Scale Estimates

After deleting the DF, t Value, Pr>|t|, and Alpha columns, abbreviating standard error as SE, and rounding all numbers to 2 decimal places, the results are shown in Table 1. The table is somewhat confusing because there are two sets of estimates labeled "Estimate" and "Mean". The column labeled "Estimate" is the summary on the link-transformed (i.e. logit) scale. The column labeled "Mean" is on the original data scale.   To convert the "Estimates (E)" to "Means (M)", one can use the formula:
$M = exp(E)/[1 + exp(E)]$. For example, for row 1, $exp(0.36)/[1 + exp(0.36)] = 0.59$. Thus, the estimated probability of a child in the Alabama control group at post measurement being properly restrained is 59% (95% CI is 37% to 78%). Similarly, E can be estimated from M as: $E = ln(M/(1-M))$. $E=ln(0.59/0.41)=0.36$ for row 1.

Table 1: Post*test*state Least square means

state	post	test	Logit scale estimates $E=ln(M/[1-M])$				Data scale Estimates (ilink) $M=exp(E)/[1+exp(E)]$			
			Estimate (E)	SE	Lower	Upper	Mean (M)	SE Mean	Lower Mean	Upper Mean
AL	Post	C	0.36	0.45	-0.53	1.25	0.59	0.11	0.37	0.78
AR	Post	C	-1.35	0.32	-1.98	-0.72	0.21	0.05	0.12	0.33
IL	Post	C	0.53	0.42	-0.30	1.35	0.63	0.10	0.43	0.79
IN	Post	C	0.21	0.47	-0.71	1.12	0.55	0.12	0.33	0.75
AL	Post	I	0.50	0.46	-0.40	1.40	0.62	0.11	0.40	0.80
AR	Post	I	-0.32	0.26	-0.83	0.19	0.42	0.06	0.30	0.55
IL	Post	I	0.58	0.44	-0.28	1.45	0.64	0.10	0.43	0.81

			Logit scale estimates E=ln(M/[1-M])				Data scale Estimates (ilink) M=exp(E)/[1+exp(E)]			
IN	Post	I	0.05	0.46	-0.85	0.94	0.51	0.11	0.30	0.72
AL	Pre	C	0.41	0.46	-0.49	1.31	0.60	0.11	0.38	0.79
AR	Pre	C	-0.99	0.27	-1.52	-0.47	0.27	0.05	0.18	0.39
IL	Pre	C	0.57	0.40	-0.22	1.35	0.64	0.09	0.45	0.79
IN	Pre	C	0.03	0.45	-0.85	0.91	0.51	0.11	0.30	0.71
AL	Pre	I	-0.19	0.46	-1.09	0.72	0.45	0.11	0.25	0.67
AR	Pre	I	-0.98	0.25	-1.48	-0.48	0.27	0.05	0.19	0.38
IL	Pre	I	0.25	0.42	-0.58	1.08	0.56	0.10	0.36	0.75
IN	Pre	I	0.40	0.45	-0.48	1.27	0.60	0.11	0.38	0.78

The GLIMMIX procedure can create a plot of the estimated proportions/probabilities which is invoked by the plots=(meanplot(ilink cl)) option on the PROC GLIMMIX statement. The ILINK command requests estimates on the original data scale.   The CL option requests that 95% (default alpha) confidence intervals be plotted. An example is shown in Figure 1.

Figure 1: Predicted probabilities of adherence to AAP standards by Time, Treatment group, and State.
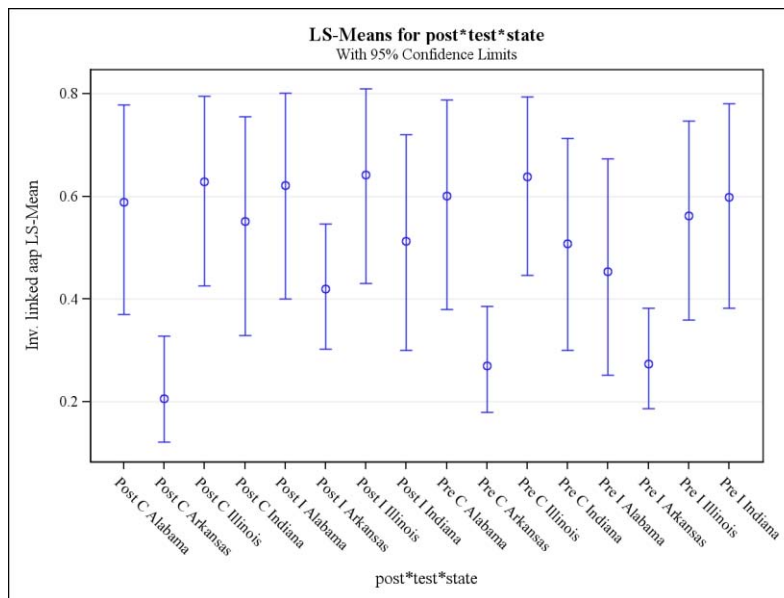
Objective 2: Within each state and treatment group (control, intervention), did the probability of meeting AAP change from pre- to post- assessment?

This question can be addressed using the LSMEANS statement with the SLICE option.

```
lsmeans state*post*test /slice=state*test;
```

For each combination of state and treatment, the hypothesis that no change occurs from pre- to post- assessment is tested.   The p-values for the tests are shown in Table 2.

Table 2: Tests of Effect Slices for post*test*state Sliced by test*state

State	test	Num DF	Den DF	F Value	Pr > F
Alabama	C	1	1745	0.03	0.8733
Arkansas	C	1	1745	1.47	0.226
Illinois	C	1	1745	0.02	0.9005
Indiana	C	1	1745	0.32	0.5723
Alabama	I	1	1745	4.71	0.0301
Arkansas	I	1	1745	9.42	0.0022
Illinois	I	1	1745	0.81	0.368
Indiana	I	1	1745	1.46	0.227

However, the slice option doesn't produce the corresponding odds ratio estimates. For simple comparisons, we can take advantage of the powerful options on the `lsmeans` statement. In order to get the odds ratios, we use an `lsmeans` statement with the `diffs` and `or` options to get odds ratios and the **`cl`** option to get 95% confidence intervals for the odds ratios. An odds ratio is a generalization of a comparison between two groups.

```
    lsmeans state*post*test / diffs or cl;
```

This `lsmeans` statement will yield all possible comparisons by `state`, `post`, and `test`. Given that this process will create a lot of extraneous output, it is helpful to save the estimates to an ODS data set and retain only the comparisons of interest.   In the following code, we select comparisons of interest: (1) comparisons within same state (or overall – state is missing), (2) comparisons of post-vs-pre within the same test group, or (3) comparisons between different test groups at the same time. If desired, we could also write SAS code using PROC MULTTEST to adjust for multiple comparisons (not shown).

```
    ods output diffs=d1;
proc glimmix data=pass2;
    class  POST TEST state ;
    model AAP=POST|TEST|state / solution dist=binary;
    random intercept   /sub=community;
    lsmeans post*test*state post*test / diff cl or ilink;
run;

data d1r;  set d1;
    **retain comparisons on same state;
    if state = _state;
    **retain comparisons within the same test group (pre-vs-post), or
    between test and controls at the same time;
    if test=_test | post= _post;
    drop estimate stderr df tvalue alpha effect lower upper;
run;
```

One disadvantage to using LSMEANS to calculate odds ratios, is that the reference group for the odds ratios may need to be inverted. For example, if odds of AAP in Arkansas for pre-vs-post, we might need to invert the odds ratio to obtain post-vs-pre. A simple data step can be used to invert the ratios as follows: ORi = 1/OR, L95ORi =1/U95OR, and U95ORi = 1/L95OR.

CALCULATIONS REQUIRING ESTIMATE STATEMENTS

For complex comparisons, it may be necessary to write `estimate` statements to obtain the estimates of interest. For contrasts involving interactions, it can be confusing to get the order of coefficients correctly entered. By including the `solution` option on the model statement, we obtain a table of model parameters in the proper order. Table 3 contains the columns Effect, state, post, and test from the estimates table.   Note that the order of variables on the CLASS statement determines the parameterization of model interactions. On the class statement, the order was post, test, and state.

Suppose we want to evaluate the change in Alabama intervention group (post-vs-pre) adjusted by the change in Alabama controls (post-vs-pre).   For the comparison of interest, we need a linear combination of 4 estimates:
1. Alabama controls at post assessment (L1).
2. Alabama controls at pre assessment (L2).
3. Alabama interventions at post assessment (L3).
4. Alabama interventions at pre assessment (L4).

Coefficients for each of these 4 estimates are listed in Table 3 as well as columns for the differences (1) D1 = L1-L2 (Alabama controls, post-vs-pre), (2) D2=L3-L4 (Alabama intervention, post-vs-pre), and (3) D2-D1 (the change in Alabama intervention group (post-vs-pre) minus the change in Alabama controls. These coefficients can be used to generate the appropriate Estimate statements.

Table 3: Solutions for Fixed Effects – Setting up contrast for AL intervention and controls, post-vs-pre

Effect	state	post	test	AL Controls (C)			AL interventions (I),			AL Change in Intervention group minus Change in Controls
				Post (L1)	Pre (L2)	Diff1= L1-L2	Post (L3)	Pre (L4)	Diff2= L3-L4	Diff2-Diff1 = L3-L4-L1+L2
Intercept				1	1	0	1	1	0	0
post		Post		1	0	1	1	0	1	0
post		Pre		0	1	-1	0	1	-1	0
test			C	1	1	0	0	0	0	0
test			I	0	0	0	1	1	0	0
post*test		Post	C	1	0	1	0	0	0	-1
post*test		Post	I	0	0	0	1	0	1	1
post*test		Pre	C	0	1	-1	0	0	0	1
post*test		Pre	I	0	0	0	0	1	-1	-1
state	AL			1	1	0	1	1	0	0
state	AR			0	0	0	0	0	0	0
state	IL			0	0	0	0	0	0	0
state	IN			0	0	0	0	0	0	0
post*state	AL	Post		1	0	1	1	0	1	0
post*state	AR	Post		0	0	0	0	0	0	0
post*state	IL	Post		0	0	0	0	0	0	0
post*state	IN	Post		0	0	0	0	0	0	0
post*state	AL	Pre		0	1	-1	0	1	-1	0
post*state	AR	Pre		0	0	0	0	0	0	0
post*state	IL	Pre		0	0	0	0	0	0	0
post*state	IN	Pre		0	0	0	0	0	0	0
test*state	AL		C	1	1	0	0	0	0	0

Effect	state	post	test	AL Controls (C)			AL interventions (I),			AL Change in Intervention group minus Change in Controls
				Post (L1)	Pre (L2)	Diff1= L1-L2	Post (L3)	Pre (L4)	Diff2= L3-L4	Diff2-Diff1 = L3-L4-L1+L2
test*state	AR		C	0	0	0	0	0	0	0
test*state	IL		C	0	0	0	0	0	0	0
test*state	IN		C	0	0	0	0	0	0	0
test*state	AL		I	0	0	0	1	1	0	0
test*state	AR		I	0	0	0	0	0	0	0
test*state	IL		I	0	0	0	0	0	0	0
test*state	IN		I	0	0	0	0	0	0	0
post*test*state	AL	Post	C	1	0	1	0	0	0	-1
post*test*state	AR	Post	C	0	0	0	0	0	0	0
post*test*state	IL	Post	C	0	0	0	0	0	0	0
post*test*state	IN	Post	C	0	0	0	0	0	0	0
post*test*state	AL	Post	I	0	0	0	1	0	1	1
post*test*state	AR	Post	I	0	0	0	0	0	0	0
post*test*state	IL	Post	I	0	0	0	0	0	0	0
post*test*state	IN	Post	I	0	0	0	0	0	0	0
post*test*state	AL	Pre	C	0	1	-1	0	0	0	1
post*test*state	AR	Pre	C	0	0	0	0	0	0	0
post*test*state	IL	Pre	C	0	0	0	0	0	0	0
post*test*state	IN	Pre	C	0	0	0	0	0	0	0
post*test*state	AL	Pre	I	0	0	0	0	1	-1	-1
post*test*state	AR	Pre	I	0	0	0	0	0	0	0
post*test*state	IL	Pre	I	0	0	0	0	0	0	0

Objective 3: Estimate statements

**Compute change in Alabama controls, post-vs-pre;

```
estimate 'AL Control - post-vs-pre' post 1 -1 post*test 1 0 -1 0 post*state 1 0 0 0
-1 0 0 0 post*test*state 1 0 0 0  0 0 0 0  -1 0 0 0   0 0 0 0 / oddsratio cl;
```

**Compute change in Alabama interventions, post-vs-pre;

```
estimate 'AL Intervention - post-vs-pre' post 1 -1 post*test 0 1 0 -1 post*state 1
0 0 0 -1 0 0 0 post*test*state 0 0 0 0  1 0 0 0   0 0 0 0   -1 0 0 0 /  oddsratio
cl;
```

** AL - Compute change in interventions (post vs pre) adjusted for change in controls;

```
Estimate 'AL change in interventions-vs-change in controls' post*test -1 1 1 -1
post*test*state -1 0 0 0 1 0 0 0 1 0 0 0 -1 0 0 0 / oddsratio cl ilink;
```

The resulting estimates can be found in Table 4. Similar code can be used to calculate results for other states.

TABLE 4: OBJECTIVE 3 ESTIMATES TABLE (AFTER DELETING DF, T VALUE, AND ALPHA COLUMNS, AND ROUNDING ESTIMATES TO 2 DECIMAL PLASES.

Label	Estimate	SE	Pr>|t|	Lower	Upper	Exp Est*	Exp Low*	Exp Up*
AL Control - post-vs-pre	-0.05	0.31	0.8733	-0.65	0.56	0.95	0.52	1.74
AL Intervention - post-vs-pre	0.68	0.31	0.0301	0.07	1.30	1.98	1.07	3.67
AL change in interventions-vs-change in controls	0.73	0.44	0.0971	-0.13	1.60	2.08	0.88	4.93

*Standard Error was abbreviated as SE. Exp Est =Exponentiated estimate (i.e. odds ratio), Exp Low= Exponentianted Lower (L95% Odds ratio), and Exp Up = Exponentiated upper (U95% Odds ratio). Again PROC GLIMMIX provides two sets of estimates: "Estimates, Lower, and Upper" and "Exponentiated estimate, Exponentiated Lower, and Exponentiated Upper." The First set corresponds to the logit scale. The exponentiated estimates correspond to the Odds ratios, and the 95% confidence intervals for the odds ratios.

OBJECTIVE 4: OVERALL CALCULATIONS – COMBINING STATE RESULTS

In order to evaluate the overall intervention, we want to pool the results from the individual states into overall estimates. For example, we might want to estimate the odds of AAP within the control (or intervention group) post versus pre. We could use either an `lsmeans` statement or `estimate` statements to obtain the odds ratios, 95% confidence intervals, and p-value:

```
lsmeans post*test / diff or cl;
estimate "Intervention, post-pre" post 1 -1 post*test 0 1 0 -1 /cl oddsratio ;
estimate "Control, post-pre"  post 1 -1 post*test 1 0 -1 0 / cl oddsratio;
```
.

However, we need to consider how SAS is handling the effect of `state` given that our model contains the 3-way interaction `post*test*state`. We can use the "e" option on the estimate statement to obtain a list of the coefficients SAS is using in the estimates.

```
estimate "Intervent, post-pre" post 1 -1 post*test 0 1 0 -1 /e ;
```

The coefficients for these comparisons are listed in Table 5.

Table 5: Coefficients for Fixed Effects (averaging over states), (1) controls, post-vs-pre, (2) intervention group, post vs pre.

Effect	state	post	test	Controls in all 4 states, post vs pre Post (L5)	Pre (L6)	Diff (L5-L6)	Interventions in all 4 states, post vs pre Post (L7)	Pre (L8)	Diff: L7-L8
Intercept				1	1	0	1	1	0
post		Post		1	0	1	1	0	1
post		Pre		0	1	-1	0	1	-1
test			C	1	1	0	0	0	0
test			I	0	0	0	1	1	0
post*test		Post	C	1	0	1	0	0	0
post*test		Post	I	0	0	0	1	0	1
post*test		Pre	C	0	1	-1	0	0	0
post*test		Pre	I	0	0	0	0	1	-1
state	AL			0.25	0.25	0	0.25	0.25	0
state	AR			0.25	0.25	0	0.25	0.25	0
state	IL			0.25	0.25	0	0.25	0.25	0
state	IN			0.25	0.25	0	0.25	0.25	0
post*state	AL	Post		0.25	0	0.25	0.25	0	0.25
post*state	AR	Post		0.25	0	0.25	0.25	0	0.25
post*state	IL	Post		0.25	0	0.25	0.25	0	0.25
post*state	IN	Post		0.25	0	0.25	0.25	0	0.25
post*state	AL	Pre		0	0.25	-0.25	0	0.25	-0.25
post*state	AR	Pre		0	0.25	-0.25	0	0.25	-0.25
post*state	IL	Pre		0	0.25	-0.25	0	0.25	-0.25
post*state	IN	Pre		0	0.25	-0.25	0	0.25	-0.25
test*state	AL		C	0.25	0.25	0	0	0	0
test*state	AR		C	0.25	0.25	0	0	0	0

Effect	state	post	test	Controls in all 4 states, post vs pre				Interventions in all 4 states, post vs pre		
				Post (L5)	Pre (L6)	Diff (L5-L6)		Post (L7)	Pre (L8)	Diff: L7-L8
test*state	IL		C	0.25	0.25	0		0	0	0
test*state	IN		C	0.25	0.25	0		0	0	0
test*state	AL		I	0	0	0		0.25	0.25	0
test*state	AR		I	0	0	0		0.25	0.25	0
test*state	IL		I	0	0	0		0.25	0.25	0
test*state	IN		I	0	0	0		0.25	0.25	0
post*test*state	AL	Post	C	0.25	0	0.25		0	0	0
post*test*state	AR	Post	C	0.25	0	0.25		0	0	0
post*test*state	IL	Post	C	0.25	0	0.25		0	0	0
post*test*state	IN	Post	C	0.25	0	0.25		0	0	0
post*test*state	AL	Post	I	0	0	0		0.25	0	0.25
post*test*state	AR	Post	I	0	0	0		0.25	0	0.25
post*test*state	IL	Post	I	0	0	0		0.25	0	0.25
post*test*state	IN	Post	I	0	0	0		0.25	0	0.25
post*test*state	AL	Pre	C	0	0.25	-0.25		0	0	0
post*test*state	AR	Pre	C	0	0.25	-0.25		0	0	0
post*test*state	IL	Pre	C	0	0.25	-0.25		0	0	0
post*test*state	IN	Pre	C	0	0.25	-0.25		0	0	0
post*test*state	AL	Pre	I	0	0	0		0	0.25	-0.25
post*test*state	AR	Pre	I	0	0	0		0	0.25	-0.25
post*test*state	IL	Pre	I	0	0	0		0	0.25	-0.25
post*test*state	IN	Pre	I	0	0	0		0	0.25	-0.25

When we leave out an effect contained in an interaction such as state, SAS will still produce the tests and estimates assuming that we have a balanced design and weight each state equally. There are 4 states, so each state has a weight of 0.25. As long as we want to weight each state equally, we can use:

```
estimate "Intervention, post-pre" post 1 -1 post*test 0 1 0 -1 /cl oddsratio ;
estimate "Control, post-pre" post 1 -1 post*test 1 0 -1 0  / cl oddsratio ;
estimate "Post, Int-control"  test -1 1 post*test -1 1 0 0 / cl oddsratio;
estimate "Pre, Int-control"   test -1 1 post*test  0 0 -1 1 / cl oddsratio;
```

When we submit the ESTIMATE statement:

```
estimate "Intervention, post-pre" post 1 -1 post*test 0 1 0 -1 /cl oddsratio ;
```

SAS assumes that we want to weight each state equally and actually estimates:

```
estimate "Intervent, post-pre" post 1 -1 post*test 0 1 0 -1 post*state  0.25 0.25
0.25 0.25 -0.25 -0.25 -0.25 -0.25 Post*test*state 0 0 0 0 0.25 0.25 0.25 0.25 0 0 0
0 -0.25 -0.25 -0.25 -0.25 /cl oddsratio ;
```

The resulting estimates are in Table 6.

Table 6: Objective 4 Estimates table + extras (after deleting df, t Value, and Alpha columns, and rounding estimates to 2 decimal places.

Label	Estimate	SE	Pr>|t|	Lower	Upper	Exp Est*	Exp Low*	Exp Up*
Intervention, post-pre	0.33	0.15	0.0285	0.03	0.63	1.39	1.04	1.87
Control, post-pre	-0.07	0.15	0.6592	-0.37	0.23	0.93	0.69	1.26
Post, Int-control	0.27	0.27	0.3179	-0.26	0.79	1.30	0.77	2.19
Pre, Int-control	-0.13	0.25	0.5991	-0.63	0.36	0.88	0.53	1.44

*Standard Error was abbreviated as SE. Exp Est =Exponentiated estimate, Exp Low= Exponentianted Lower, and Exp Up = Exponentiated upper. Again PROC GLIMMIX provides two sets of estimates: "Estimates, Lower, and Upper" and "Exponentiated estimate, Exponentiated Lower, and Exponentiated Upper." The First set corresponds to the logit scale. The exponentiated estimates correspond to the Odds ratios, and the 95% confidence intervals for the odds ratios.

SUPPOSE YOU WANT TO WEIGHT THE STATES

As discussed in David Pasta's paper, there may be times that you want to weight levels of a factor unequally. If we have 23 interventions with 4 in Alabama, 11 in Arkansas, 4 in Illinois, and 4 in Indiana.   Given that nearly half of the interventions occurred in Arkansas, we might want to give Arkansas more influence in the overall calculations than the other states. One might decide to weight the estimates based on the number of interventions in each state. Thus, we would use 4/23 (AL), 11/23 (AR), 4/23 (IL), and 4/23 (IN) as the weights for overall calculations rather than the default calculation weighting each equally (0.25, 0.25, 0.25, 0.25).   Rather than use 0.47826087 as the weight for Arkansas (i.e. 11/23), we can use a weight of 11 with the divisor option of 23 on the estimate statement.

So instead of using the overall estimate of the effect of "intervention group, post versus pre" which assumes each state is equal (i.e.):

```
estimate "simple" post 1 -1 post*test 0 1 0 -1 /cl oddsratio ;
```
which is interpreted as:
```
estimate "detailed" post 1 -1 post*test 0 1 0 -1 post*state  0.25 0.25 0.25 0.25 -
0.25 -0.25 -0.25 -0.25 Post*test*state 0 0 0 0 0.25 0.25 0.25 0.25 0 0 0 0 -0.25 -
0.25 -0.25 -0.25 /cl oddsratio ;
```
we could weight each state by the number of interventions as follows:
```
estimate "reweighted" post 23 -23 post*test 0 23 0 -23
post*state 4  11  4  4 -4 -11 -4 -4 Post*test*state 0 0 0 0 4 11 4 4 0 0 0 0 -4 -11
-4 -4 /cl oddsratio divisor=23;
```
Note that the divisor=23 indicates that each coefficient in the `estimate` statement will be divided by 23. The resulting estimates are contained in table 7.

Table 7 – Estimates of change in the intervention group – simple specification, detailed specification, and reweighting states

Label	Est	SE	p-value	Lower	Upper	Exp Est	Exp Low	Exp Up
Simple	0.33	0.15	0.0285	0.03	0.63	1.39	1.04	1.87
detailed	0.33	0.15	0.0285	0.03	0.63	1.39	1.04	1.87
reweighted	0.43	0.14	0.0025	0.15	0.71	1.54	1.16	2.03

*Standard Error was abbreviated as SE. Exp Est =Exponentiated estimate (i.e. odds ratio), Exp Low= Exponentianted Lower (L95% Odds ratio), and Exp Up = Exponentiated upper (U95% Odds ratio). Again PROC GLIMMIX provides two sets of estimates: "Estimates, Lower, and Upper" and "Exponentiated estimate, Exponentiated Lower, and Exponentiated Upper." The First set corresponds to the logit scale. The exponentiated estimates correspond to the Odds ratios, and the 95% confidence intervals for the odds ratios.

CONCLUSION

We have demonstrated many useful options for post-hoc estimation within the SAS® GLIMMIX procedure. When the response is binary, PROC GLIMMIX can produce estimates that are link transformed, proportions, and odds ratios, depending on the options selected. Our paper concentrated on contrasts among fixed effects. PROC GLIMMIX can also be used to make inferences involving random effects as well, but that is beyond the scope of this paper.

REFERENCES

Hayes RJ and Moulton LH. Cluster Randomised Trials. Chapman & Hall/CRC, Boca Raton, FL; 2009.

Murray DM, Van Horn ML, Hawkins JD and Arthur MW. Analysis strategies for a community trial to reduce adolescent ATOD use: A comparison of random coefficient and ANOVA/ANCOVA models. *Contemporary Clinical Trials*, 2006: 27: 188-206.

Pasta, David J. (2010), "Practicalities of Using ESTIMATE and CONTRAST Statements," Proceedings of the SAS Global Forum 2010, Paper 269-2010.
http://support.sas.com/resources/papers/proceedings10/269-2010.pdf

ACKNOWLEDGMENTS

CONTACT INFORMATION

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

Name: Jeff Gossett
Enterprise: UAMS Pediatrics
E-mail: gossettjeffreym@uams.edu