SAS Statistics Research and Applications
Paper #2022-03

Computing Predictive Margins for Generalized Linear Models with PROC GLIMMIX

Shiying Wu, SAS Institute Inc., Cary, NC

Using the estimated regression coefficient to interpret a covariate (or regressor) effect might be the best approach only in the context of fitting a linear regression model. A more practical and universal way to assess and directly interpret the impact of a covariate on the outcome response of any regression-type model is through the computation of predictive margins. A predictive margin is the average predicted response from a model when the designated covariate is set to a specific level for all sample observations. The difference between the predictive margins at two different levels (say, a treatment level and a baseline level) of the covariate represents a marginal effect. Marginal effects are most useful in providing interpretable results for any type of regression model, whether they are linear or nonlinear and with or without interaction effects. This paper introduces the concepts of predictive margins and marginal effects, explains variations of these concepts, illustrates how you can compute these quantities by applying the popular GLIMMIX procedure in SAS/STAT® software to numerical examples, and discusses how you can use predictive margins to interpret effects and make statistical inferences in practical applications.

Introduction

Estimated regression coefficients have long been taught as the preferred way to understand a statistical model. They are viewed as a crucial component in connecting the covariates with the outcome of interests. However, this way of thinking is hitting its limit, as practitioners attempt to model more complex relationships by relying on nonlinear models. The coefficients, along with information such as standard errors and confidence intervals, can no longer be used to enable direct and understandable interpretation, much less to answer questions such as what is the precise impact a covariate has on the outcome while accounting for population variabilities in other covariates. For example, in a linear regression model with sex as a covariate, the response from male could be used as the baseline, and the female effect on the response could be represented by its regression coefficient. That is, if the estimated regression coefficient for female is 5, then, everything else being equal, the predicted response of a female is 5 units higher than that of a male.

However, this is not true in a logistic regression model that has the same covariates. The logistic function converts the linear combination of the regression coefficients to a response probability, and a regression coefficient of 5 no longer represents the quantitative change in response probability, which is always between 0 and 1. To see the effect of a designated covariate on the scale of the response, the average predicted responses at some specific levels of the covariate are calculated. These averages are called predictive margins. The difference between the predictive margins at two specific levels of the covariate represents a marginal effect. In particular, the marginal effect of a specific level (against its baseline level) of the designated covariate is the nonlinear equivalent of the covariate effect in the linear model. For example, in a clinical trial of a new COVID-19 vaccine, it would be particularly useful to know the reduction in predicted infection rate due to vaccination against the baseline level. Such a reduction rate provides a more direct interpretation than a statistic like a regression coefficient estimate of a nonlinear model. The reductions are specific marginal effects that can be computed by the MARGINS statement in PROC GLIMMIX. The standard errors of predictive margins and their differences are computed using the delta method. Note that the model might contain interaction terms that involve the designated covariates. In this case, the marginal effects represent the average "total" change in response, including the main effects and the interactions.

Predictive margins are often used to quantify the effect of a treatment for hypothetical and possibly counterfactual cases. For example, if you want to quantify the difference in predicted response between treatment and control groups when all patients are of a specific age, you could compute the marginal effect between treatment and control while setting all patients' ages to that age, regardless of their observed ages.

In the next section, we provide more details about nonlinear models and predictive margins. The section after that introduces the MARGINS statement in PROC GLIMMIX and demonstrates how it is used to obtain estimated predictive margins and marginal effects. Examples are used in these sections to illustrate the concept, the computation, and the applications of the MARGINS statement. In the last section, we provide some concluding remarks.

Example 1: Predictive Margins

Consider a clinical study that investigates the performance of three treatments in terms of their average response rate. Patients are randomly assigned to treatment A, B, or C. In addition, each patient is also measured for smoking status (heavy smoker, light smoker, or nonsmoker).

The problem can be formulated as follows.

For patient j, if π_j is the response probability, a logistic regression model for π_j is

$$\text{logit}(\pi_j) = \beta_0 + \beta_T + \beta_S + \beta_{TS} + \beta_1 \text{Age}_j \tag{1}$$

where β_T, β_S, and β_{TS} are the effects of treatment (trt), smoking, and the trt*smoking interaction, respectively. To be more specific, $\beta_T = \beta_A$ if patient j receives treatment A, $\beta_T = \beta_B$ if patient j receives treatment B, and so on. Similarly, $\beta_S = \beta_H$ if patient j is a heavy smoker, $\beta_S = \beta_L$ if patient j is a light smoker, and so on. Finally, β_{TS} is one of the nine combinations of the treatment and smoking values: β_{AH}, β_{AL}, β_{AN}, β_{BH}, β_{BL}, β_{BN}, β_{CH}, β_{CL}, and β_{CN}. The right-hand side is also called the linear predictor for patient j:

$$\eta_j = \beta_0 + \beta_T + \beta_S + \beta_{TS} + \beta_1 \text{Age}_j$$

Given this model, the predictive margin for treatment A is

$$\hat{\pi}_A = \frac{1}{n}\sum_{j=1}^{n} g^{-1}(\hat{\beta}_0 + \hat{\beta}_A + \hat{\beta}_S + \hat{\beta}_{AS} + \hat{\beta}_1 \mathrm{Age}_j) \tag{2}$$

where $g^{-1}(\cdot)$ is the inverse logistic link function. Note that whereas the treatment effect $\hat{\beta}_T = \hat{\beta}_A$ for all individuals, $\hat{\beta}_S$ would take different values for different individuals; that is, $\hat{\beta}_S = \hat{\beta}_H$ if patient j is a heavy smoker, $\hat{\beta}_S = \hat{\beta}_L$ if the patient j is a light smoker, and so on. Similarly, the interaction term $\hat{\beta}_{AS}$ takes the three values that involve treatment A ($\hat{\beta}_{AH}$, $\hat{\beta}_{AL}$, and $\hat{\beta}_{AN}$), depending on the smoking status of the observation. An interpretation of the predictive margin that is computed in (2) is that it is the average predicted response rate if all patients were to receive treatment A. Because in reality not all patients would receive treatment A, the predictive margin can be viewed as a counterfactual concept, which has proven to be a useful construct in the field of causal analysis.

Similarly, the predictive margin for treatment B is

$$\hat{\pi}_B = \frac{1}{n}\sum_{j=1}^{n} g^{-1}(\hat{\beta}_0 + \hat{\beta}_B + \hat{\beta}_S + \hat{\beta}_{BS} + \hat{\beta}_1 \mathrm{Age}_j) \tag{3}$$

The predictive margin that is computed in this equation is the average predicated response rate if all patients were to receive treatment B.

To compare the average response rates, you can compute the marginal effect

$$\hat{\pi}_A - \hat{\pi}_B \tag{4}$$

Now let's compute the variance of $\hat{\pi}_A$ by using the delta method. Denote $\boldsymbol{\beta} = (\beta_0, \beta_A, \beta_B, \beta_C, \beta_H, \ldots, \beta_{CN}, \beta_1)'$ and $\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \hat{\beta}_A, \hat{\beta}_B, \hat{\beta}_C, \hat{\beta}_H, \ldots, \hat{\beta}_{CN}, \hat{\beta}_1)'$, and let its estimated covariance be $\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\beta}})$. Then the variance of $\hat{\pi}_A$ is estimated by

$$\hat{V}(\hat{\pi}_A) = \frac{1}{n^2}\sum_{j=1}^{n} \Delta_j \widehat{\mathrm{Cov}}(\hat{\boldsymbol{\beta}}) \Delta_j'$$

where Δ_j contains the partial derivatives of $g^{-1}(\beta_0 + \beta_A + \beta_S + \beta_{AS} + \beta_1 \mathrm{Age}_j)$ with respect to $\boldsymbol{\beta}$ evaluated at $\hat{\boldsymbol{\beta}}$; $\beta_S = \beta_H$ if patient j is a heavy smoker; $\beta_S = \beta_L$ if patient j is a light smoker; and so on. Similarly, the interaction term β_{AS} takes the three values that involve treatment A (β_{AH}, β_{AL}, and β_{AN}), depending on the smoking status of the observation.

More generally, in a generalized linear model, $g(\cdot)$ could be any link function, and there could be any number of classification and continuous covariates with interactions of any order. The margins are computed the same way. That is, you set the designated covariate to the given level while setting all other covariates to the observed values, plug in the covariates and the estimated regression coefficients $\hat{\beta}$ to compute the estimated linear predictor $\hat{\eta}_j$, and then average $g^{-1}(\hat{\eta}_j)$ over all j. The variances of the margins are also computed the same way by applying the delta method.

The computation of predictive margins appears to be similar to the computation of least squares means (LS-means), but with two important differences: (1) Predictive margins are constructed on the mean scale, whereas LS-means are constructed on the linked scale; and (2) predictive margins account for covariate imbalances, whereas LS-means estimate the marginal means over a balanced population.

MARGINS Statement in PROC GLIMMIX

In this section, we show how to obtain predictive margins in SAS® via the MARGINS statement, in an example and in more general cases.

The MARGINS statement in PROC GLIMMIX computes predictive margins of fixed effects in a multilevel model. In this paper, we focus on generalized linear models, which are a special class of multilevel models. Predictive margins can be computed for any fixed effects of classification variables in the MODEL statement. The syntax of the MARGINS statement is as follows:

```
MARGINS fixed-effects < / options > ;
```

You can specify multiple effects in one MARGINS statement or in multiple MARGINS statements, and all MARGINS statements must appear after the MODEL statement. PROC GLIMMIX constructs an approximate t test to test the null hypothesis that the corresponding population parameter equals zero. By default, the denominator degrees of freedom for this test are the same as those displayed for the effect in the "Type III Tests of Fixed Effects" table.

Table 1 summarizes the options of the MARGINS statement.

Table 1: MARGINS Statement Options

Option	Description
Construction of Predictive Margins	
AT	Specifies the covariate value to use in computing predictive margins
DIFF	Requests differences of predictive margins
SLICEBY=	Partitions F tests
SLICEDIFF=	Requests differences of sliced predictive margins and determines the type of differences
Degrees of Freedom and p-Values	
ADJUST=	Specifies the method of multiple comparison adjustment of predictive margin differences
ALPHA=α	Specifies the confidence level $(1-\alpha)$
DF=	Assigns a specific value to degrees of freedom for tests and confidence limits
STEPDOWN	Adjusts multiple comparison p-values further in a step-down fashion
Statistical Output	
CL	Constructs confidence limits for predictive margins and predictive margin differences

Now let's use the previous example to demonstrate how you can use PROC GLIMMIX to compute the predictive margins. As in that example, trt and Smoking are CLASS statement variables and Age is a continuous variable in the model for predicting a binary outcome. You can compute predictive margins in three different ways:

- Compute the prediction at the observed values of other covariates:

```
proc glimmix;
   class trt smoking;
   model response = trt smoking trt*smoking age / dist=binary;
   margins trt/ cl;
run;
```

  By default, the MARGINS statement sets Smoking and Age to their observed values for each observation and computes the average response of the outcome at each level of the CLASS variable trt. The CL option computes the confident limits of the margins.

- Compute the prediction at the means of other covariates:

```
proc glimmix;
   class trt smoking;
   model response = trt smoking trt*smoking age / dist=binary;
   margins trt /cl at means;
run;
```

  The MARGINS statement sets the continuous variable Age to its mean and the CLASS variable Smoking to its observed value for each observation, and it computes the average response of the outcome at each level of the CLASS variable trt. As a rule, the various AT options apply only to continuous covariates, not to categorical covariates, which are always computed at the observed values of individuals.

- Compute the predictions at specific values of other covariates:

```
proc glimmix;
   class trt smoking;
   model response = trt smoking trt*smoking age / dist=binary;
   margins trt /cl at age=50;
run;
```

  The MARGINS statement sets Age to 50 for each observation and computes the average response of the outcome at each level of the CLASS variable trt. Again, the CLASS variable Smoking is set to its observed value.

Corresponding to the three ways of computing predictive margins, there are three ways that you can compute marginal effects. The marginal effect for categorical variables shows how the probability of an observed event, such as PROB(OUTCOME=1), changes as a categorical variable changes from the reference level to a treatment level, after controlling for the other variables in the model. Hence, marginal effects are computed as differences of predictive margins. The following examples use the DIFF option to compute the marginal effect.

- Compute the marginal effect at the observed values of other covariates:

```
proc glimmix;
   class trt smoking;
   model response = trt smoking trt*smoking age / dist=binary;
   margins trt / cl diff;
run;
```

- Compute the marginal effect at the means of continuous covariates:

```
proc glimmix;
   class trt smoking;
   model response = trt smoking trt*smoking age / dist=binary;
   margins trt /cl diff at means;
run;
```

- Compute the marginal effect at specific values of continuous covariates:

```
proc glimmix;
   class trt smoking;
   model response = trt smoking trt*smoking age / dist=binary;
   margins trt / cl diff at age=50;
run;
```

Note that using the AT option to specify the covariate values at their means or specific values, rather than using their observed values, might be controversial in some cases. Some researchers have doubts about the appropriateness of this usage, because it is possible that no one in the population takes the mean or specified values. In other words, the computed margins or marginal differences might not correspond to a meaningful population of interest. In addition, either of these methods with fixed covariate values misses variability in effects across cases. Williams (2012) suggests specifying a range of covariate values to provide a more complete picture of how the target marginal effects differ across cases. The CL option computes both the confident limits of the margins and the marginal effects.

The MARGINS statement also performs statistical tests for marginal effects, for differences of marginal effects, and with multiple comparison adjustments for the p-values upon request. Here are some examples.

- Test the marginal effect (the effect of trt for each smoking group):

```
proc glimmix;
   class trt smoking;
   model response = trt smoking trt*smoking age / dist=binary;
   margins trt*smoking /sliceby=smoking;
run;
```

  In this example, the predictive margins of the interaction trt*Smoking are computed by averaging the predicted responses after setting (trt, Smoking) to each of their nine combinations of levels and setting Age to its observed level. The SLICEBY= option produces an F test that tests the simultaneous equality of the margins at each level of Smoking.

- Test the differences of marginal effects:

```
proc glimmix;
   class trt smoking;
   model response = trt smoking trt*smoking age / dist=binary;
   margins trt*smoking /sliceby=smoking slicediff=control;
run;
```

Here, for each level of trt, you jointly test the following hypotheses: H1, that the effects of heavy smoking versus light smoking on the outcome are the same; H2, that the effects of light smoking versus nonsmoking on the outcome are the same.

- Make multiple comparison adjustments for the p-values and confidence limits:

```
proc glimmix;
   class trt smoking;
   model response = trt smoking trt*smoking age / dist=binary;
   margins trt*smoking /sliceby=smoking slicediff=control adjust=Tukey;
run;
```

Extending the previous example, you can add the ADJUST= option to specify a multiple comparison adjustment method. The choices are BON, DUNNETT, SCHEFFE, SIDAK, SIMULATE, SMM, and TUKEY. By default, PROC GLIMMIX performs all pairwise differences. For details, see the PROC GLIMMIX documentation in SAS Institute Inc. (2022) (https://go.documentation.sas.com/doc/en/pgmsascdc/v_026/statug/statug_glimmix_syntax15.htm).

In this example, Tukey is specified as the adjustment method in the ADJUST= option. For detailed descriptions of the other choices, see the PROC GLIMMIX documentation.

Application Example

Continuing our example about a clinical study in which patients were randomly assigned to treatment A, B, or C, each patient's smoking status was also recorded (heavy smoker, light smoker, or nonsmoker), along with their age and outcome response. The data are recorded in the data set Trial, as shown in the following DATA step:

```
data trial;
   input trt$ age smoking response @@;
   datalines;
   A 15 2 0 A 28 1 0 A 24 1 0 B 54 1 0
   A 60 1 0 A 68 2 1 B 64 1 0 B 15 0 0
   A 23 2 0 A 33 1 0 B 15 2 0 B 76 1 1
   A 30 1 0 A 73 0 1 B 39 0 1 B 48 1 0
   A 15 0 1 A 34 0 0 B 34 1 0 B 17 1 0
   A 15 1 0 A 68 0 1 A 74 1 0 A 78 1 1
   A 49 1 0 B 67 1 0 A 15 1 0 A 41 0 0
   B 53 1 0 B 62 0 1 A 25 2 0 A 31 0 0
   B 15 1 1 B 28 1 0 A 22 1 0 A 15 1 0
   B 27 1 0 B 45 1 0 A 68 0 1 B 77 1 0
   B 56 1 0 B 24 1 0 B 33 1 0 B 21 1 0
   B 42 2 0 B 61 1 0 B 15 1 0 B 70 1 0
   B 15 1 0 B 67 0 1 B 54 0 1 B 15 2 0
   A 43 1 0 A 52 0 1 B 15 1 0 B 40 1 0
   A 49 2 0 A 59 0 1 A 21 2 0 A 72 1 1
   A 32 2 0 A 50 1 0 A 28 2 0 A 58 1 1
   A 41 1 0 A 21 2 0 A 39 0 1 A 41 1 0
```

```
   A 62 2 0 A 77 2 1 A 35 1 0 A 59 2 0
   A 70 1 0 B 79 1 1 B 73 0 1 B 15 0 0
   B 49 1 0 B 73 1 1 B 46 1 0 B 59 0 1
   C 73 2 0 C 78 1 1 C 20 1 0 C 66 1 1
   C 61 1 0 C 78 2 0 C 24 1 0 C 15 2 0
   C 54 2 0 C 51 1 0 C 15 0 0 C 63 0 1
   C 50 1 0 C 17 2 0 C 15 0 0 C 61 1 0
   C 15 0 0 C 20 2 0 C 73 2 1 C 35 1 0
   C 18 1 0 C 32 1 0 C 76 1 0 C 76 2 0
   C 55 1 0 C 51 0 1 C 15 2 0 C 15 2 0
   C 58 0 1 C 36 1 0 C 15 1 0 C 15 2 0
   C 15 1 0 C 21 1 0 C 47 1 0 C 77 0 1
;
run;
```

The data set contains four variables. The variable trt identifies the three treatments, A, B, and C. The variable Smoking takes the value of 2 if the patient is a heavy smoker, 1 if the patient is a light smoker, and 0 if the patient is a nonsmoker. The response variable is 1 if the patient responds to the treatment and 0 if not.

We used the model in (1) to model the output, and we want to compute the predictive margins that are defined in (2) and (3), as well as the corresponding marginal effect in (4). We also want to examine the effects of trt at each smoking level. The following statements fit the logistic regression model and compute the predictive margins and marginal effects:

```
proc glimmix data=trial;
   class trt smoking;
   model response = trt|smoking age/s dist=binary link=logit;
   margins trt/ diff;
   margins trt*smoking/ sliceby=smoking slicediff;
run;
```

The first MARGINS statement requests predictive margins for the three treatment groups. The DIFF option compares average treatment response rates, controlling for the age and smoking distributions. Table 2 shows the predictive margins for treatments A, B, and C.

Table 2: Margins

trt Margins

trt	Estimate	Standard Error	DF	t Value	Pr > |t|
A	0.7252	0.04631	106	15.66	<.0001
B	0.7671	0.03908	106	19.63	<.0001
C	0.7768	0.04972	106	15.62	<.0001

Table 3 shows the results of testing the pairwise differences of the three treatment margins.

Based on the p-value, you would conclude that at the 0.05 level, the average response rates are not significantly different among the three treatments.

The second MARGINS statement requests predictive margins for the trt*Smoking interaction. The SLICEBY= option slices the trt*Smoking interaction by smoking level. The SLICEDIFF

option then compares the sliced predictive margins for the heavy, light, and nonsmoker groups separately. Table 4 shows the predictive margins for trt and Smoking interactions.

Table 3: Pairwise Differences of the Margins

Differences of trt Margins

trt	_trt	Estimate	Standard Error	DF	t Value	Pr > |t|
A	B	-0.04189	0.06061	106	-0.69	0.4910
A	C	-0.05159	0.06791	106	-0.76	0.4491
B	C	-0.00970	0.06327	106	-0.15	0.8784

Table 4: Margins for trt*Smoking Interaction

trt*smoking Margins

trt	smoking	Estimate	Standard Error	DF	t Value	Pr > |t|
A	H	0.8029	0.09570	106	8.39	<.0001
A	L	0.8351	0.07079	106	11.80	<.0001
A	N	0.4061	0.1015	106	4.00	0.0001
B	H	1.0000	0.01578	106	63.36	<.0001
B	L	0.8601	0.05472	106	15.72	<.0001
B	N	0.3273	0.1467	106	2.23	0.0278
C	H	0.9424	0.05351	106	17.61	<.0001
C	L	0.8735	0.07039	106	12.41	<.0001
C	N	0.4158	0.1562	106	2.66	0.0090

Table 5 through Table 7 show the tests of the treatment margin differences for the three smoking groups: H (heavy smoking), L (light smoking), and N (nonsmoking).

Table 5: Tests of the Treatment Margin Differences for Heavy Smoking

Differences of trt*smoking Margins Sliced by smoking

Slice	trt	_trt	Estimate	Standard Error	DF	t Value	Pr > |t|
smoking H	A	B	-0.1971	0.09699	106	-2.03	0.0447
smoking H	A	C	-0.1395	0.1097	106	-1.27	0.2064
smoking H	B	C	0.05760	0.05579	106	1.03	0.3043

Table 6: Tests of the Treatment Margin Differences for Light Smoking

Differences of trt*smoking Margins Sliced by smoking

Slice	trt	_trt	Estimate	Standard Error	DF	t Value	Pr > |t|
smoking L	A	B	-0.02504	0.08948	106	-0.28	0.7802
smoking L	A	C	-0.03837	0.09979	106	-0.38	0.7014
smoking L	B	C	-0.01333	0.08917	106	-0.15	0.8814

Table 7: Tests of the Treatment Margin Differences for Nonsmoking

Differences of trt*smoking Margins Sliced by smoking

Slice	trt	_trt	Estimate	Standard Error	DF	t Value	Pr > |t|
smoking N	A	B	0.07878	0.1779	106	0.44	0.6588
smoking N	A	C	-0.00973	0.1863	106	-0.05	0.9584
smoking N	B	C	-0.08851	0.2143	106	-0.41	0.6804

Table 5 shows that the average response rates are significantly different between treatment A and treatment B in the heavy smoking group (p-value of 0.04), despite the fact that the average response rates are not significantly different in the light smoking group (p-value of 0.78; Table 6), the nonsmoking group (p-value of 0.66; Table 7), and the overall sample (p-value of 0.49; Table 3).

Note that the p-values of the three tests are not adjusted for multiple comparisons. To request adjusted p-values, along with adjusted confidence limits, add the ADJUST= option, as shown in the following statements:

```
proc glimmix data=trial;
   class trt smoking;
   model response = trt|smoking age/s dist=binary link=logit;
   margins trt*smoking /cl sliceby=smoking slicediff adjust=tukey;
run;
```

The adjusted results are shown in Table 8 through Table 10.

Table 8: Tests of the Treatment Margin Differences for Heavy Smoking Group with Adjustment for Multiple Comparisons

Differences of trt*smoking Margins Sliced by smoking
Adjustment for Multiple Comparisons: Tukey-Kramer

Slice	trt	_trt	Estimate	Standard Error	DF	t Value	Pr > |t|	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
smoking H	A	B	-0.1971	0.09699	106	-2.03	0.0447	0.1096	0.05	-0.3894	-0.00479	-0.4277	0.03348
smoking H	A	C	-0.1395	0.1097	106	-1.27	0.2064	0.4145	0.05	-0.3570	0.07804	-0.4003	0.1213
smoking H	B	C	0.05760	0.05579	106	1.03	0.3043	0.5582	0.05	-0.05302	0.1682	-0.07503	0.1902

Table 9: Tests of the Treatment Margin Differences for Light Smoking Group with Adjustment for Multiple Comparisons

Differences of trt*smoking Margins Sliced by smoking
Adjustment for Multiple Comparisons: Tukey-Kramer

Slice	trt	_trt	Estimate	Standard Error	DF	t Value	Pr > |t|	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
smoking L	A	B	-0.02504	0.08948	106	-0.28	0.7802	0.9578	0.05	-0.2024	0.1524	-0.2377	0.1877
smoking L	A	C	-0.03837	0.09979	106	-0.38	0.7014	0.9218	0.05	-0.2362	0.1595	-0.2756	0.1988
smoking L	B	C	-0.01333	0.08917	106	-0.15	0.8814	0.9878	0.05	-0.1901	0.1635	-0.2253	0.1986

Table 10: Tests of the Treatment Margin Differences for Nonsmoking Group with Adjustment for Multiple Comparisons

Differences of trt*smoking Margins Sliced by smoking
Adjustment for Multiple Comparisons: Tukey-Kramer

Slice	trt	_trt	Estimate	Standard Error	DF	t Value	Pr > |t|	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
smoking N	A	B	0.07878	0.1779	106	0.44	0.6588	0.8977	0.05	-0.2740	0.4315	-0.3441	0.5017
smoking N	A	C	-0.00973	0.1863	106	-0.05	0.9584	0.9985	0.05	-0.3790	0.3595	-0.4525	0.4330
smoking N	B	C	-0.08851	0.2143	106	-0.41	0.6804	0.9103	0.05	-0.5133	0.3363	-0.5978	0.4208

You can see that these output tables, in addition to displaying the usual p-value and confidence limits, also display p-values (Adj P) and confidence limits (Adj Lower, Adj Upper) that have been adjusted by multiple comparisons.

If you are interested in looking at the predictive margins and marginal effects when Age is set to its mean, you can specify the AT MEANS option in the MARGINS statement, as shown in the following statements:

```
proc glimmix data=trial;
   class trt smoking;
   model response = trt|smoking age/s dist=binary link=logit;
   margins trt/ at means cl diff ;
run;
```

The CL option in the MARGINS statement computes the confidence limits of the predictive margins, which are shown in Table 11.

Table 11: Treatment Margins at AGE=MEAN

trt Margins

trt	age	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
A	42.91	0.8085	0.04742	106	17.05	<.0001	0.05	0.7145	0.9025
B	42.91	0.8027	0.04054	106	19.80	<.0001	0.05	0.7223	0.8831
C	42.91	0.8328	0.06451	106	12.91	<.0001	0.05	0.7049	0.9607

Note that the mean value of Age is listed in the second column to indicate that the predictive margins are computed at the specific Age value.

The marginal effect, which is the difference between the predictive margins, along with its own confidence interval, is given in Table 12.

Table 12: Tests of the Treatment Margin Differences at AGE=MEAN

Differences of trt Margins

trt	_trt	age	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
A	B	42.91	0.005825	0.06194	106	0.09	0.9252	0.05	-0.1170	0.1286
A	C	42.91	-0.02427	0.07975	106	-0.30	0.7614	0.05	-0.1824	0.1338
B	C	42.91	-0.03010	0.07615	106	-0.40	0.6934	0.05	-0.1811	0.1209

Instead of using the mean value of Age, you might also compute the predictive margins at a particular age, such as 65, by using the following statements:

```
proc glimmix data=trial;
   class trt smoking;
   model response = trt|smoking age/s dist=binary link=logit;
   margins trt/ at age=65 cl;
run;
```

Table 13 shows the predictive margins at age 65.

Table 13: Tests of the Treatment Margin Differences at Age 65

trt Margins									
trt	age	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
A	65.00	0.5125	0.1118	106	4.58	<.0001	0.05	0.2908	0.7342
B	65.00	0.6353	0.07313	106	8.69	<.0001	0.05	0.4903	0.7803
C	65.00	0.6357	0.09459	106	6.72	<.0001	0.05	0.4482	0.8232

Concluding Remarks

Predictive margins are counterfactual predictions in which the target covariates are assigned to fixed levels while other covariates are kept at the observed, mean, or specific values. Marginal effects are especially useful when you want to interpret effects on the natural scale of the outcome response of interest rather than the parameter scale of the estimated model; these scales are not the same in nonlinear models. Marginal effects are used in many statistical analyses, whether it is a prospective study or an observational study. The example in this paper shows a randomized controlled trial in which response is evaluated against three different treatments. For an example of an observational study, see Lane and Nelder (1982), who look at the association between smoking status and age of natural menopause.

In summary, the MARGINS statement in PROC GLIMMIX provides a convenient way to obtain predictive margins, marginal effects, and their confidence intervals, with covariates at observed values, respective means, or user-specified values. It also performs statistical tests for marginal effects and differences of marginal effects, with multiple comparison adjustment for the p-values upon request.

References

Lane, P. W., and Nelder, J. A. (1982). "Analysis of Covariance and Standardization as Instances of Prediction." *Biometrics* 38:613–621.

SAS Institute Inc. (2022). *SAS/STAT User's Guide*. Cary, NC: SAS Institute Inc. Revised March 2022. https://documentation.sas.com/doc/en/pgmsascdc/v_026/statug/titlepage.htm.

Williams, R. (2012). "Using the Margins Command to Estimate and Interpret Adjusted Predictions and Marginal Effects." *Stata Journal* 12:308–331.

Acknowledgments

Thanks to Yiu-Fai Yung, Fang Chen, and Bob Derr for their help in preparing this paper.