Example 5: Generalized Models for Binary Longitudinal Data
(complete syntax and output available for SAS and STATA electronically)

This example comes from real data collected over 12 days in 91 nursing home patients who were hospitalized. Day 0 is the first day of hospitalization. Each day an assessment as to the patient's level of delirium was conducted by hospital staff, in which 0 = no delirium, 1 = possible delirium, and 2 = full delirium. *For the purposes of illustration*, the 1 and 2 categories were collapsed to create a binary outcome of none vs. at least some delirium, with 37.25% = none and 62.75% = at least some delirium. SPSS does not estimate generalized multilevel levels using numeric integration, so only SAS and STATA are used for these binary models. We will first examine the pattern of change across days using polynomial models, and then see if cognitive status (as measured by MMSE centered at 16, SD = 7) predicts intercept and slope differences in probability of delirium.

Model 1: Empty Logistic Two-Level Model for None (0) vs. at least Some Delirium (1)

```
TITLE1 "SAS Empty Logistic Mixed Model";
TITLE2 "Logit Link, Binomial Distribution";
PROC GLIMMIX DATA=Example5 NOCLPRINT NOITPRINT NAMELEN=100 GRADIENT METHOD=QUAD(QPOINTS=7);
      CLASS patient_ID;
      MODEL cam011 (DESCENDING) =  / SOLUTION LINK=LOGIT DIST=BINARY DDFM=BW;
      RANDOM INTERCEPT / TYPE=UN SUBJECT=patient_ID;
      ESTIMATE "Intercept" intercept 1 / ILINK;      * ILINK = inverse link to get probability;
      COVTEST "Need Random Intercept?"  0;
RUN; TITLE1; TITLE2;
```

COVTEST does a LRT for the contents of the G matrix as specified to be either 0 or . (where . means estimate it as specified). Below, the 0 value asks SAS to hold the random intercept variance to 0 and estimate that model in order to do a LRT test against our model.

```
 * STATA Empty Logistic Mixed Model,
 * Logit Link, Binomial Distribution
melogit cam011 ,  ||  PATIENT_ID:  ,
      covariance(unstructured) intpoints(7),
      estat ic, n(91),
      estat icc
```

SAS Output:

Fit Statistics

-2 Log Likelihood	627.17
AIC (smaller is better)	631.17
AICC (smaller is better)	631.19
BIC (smaller is better)	636.19

$$\text{Level-1: Logit}\left[p\left(y_{ti}=1\right)\right]=\beta_{0i}$$
$$\text{Level-2: Intercept: } \beta_{0i}=\gamma_{00}+U_{0i}$$

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Gradient	
UN(1,1)	PATIENT_ID	1.5560	0.5078	9.623E-6	Random intercept variance

Solutions for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient
Intercept	0.7410	0.1847	90	4.01	0.0001	0.000042

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
Intercept	0.7410	0.1847	90	4.01	0.0001	0.6772	0.0403

In Probability

Tests of Covariance Parameters
Based on the Likelihood

Label	DF	-2 Log Like	ChiSq	Pr > ChiSq	Note
Need Random Intercept?	1	673.50	46.34	<.0001	MI

MI: P-value based on a mixture of chi-squares.

COVTEST results: −2LL given is for a model without the random intercept variance

Model 2: Adding a Fixed Linear Slope for Days Since Hospital Admission (0=Day Hospitalized)

```
TITLE3 "SAS Add Fixed Linear Slope for Day";
PROC GLIMMIX DATA=Example5 NOCLPRINT NOITPRINT NAMELEN=100 GRADIENT METHOD=QUAD(QPOINTS=7);
      CLASS patient_ID;
      MODEL cam011 (DESCENDING) = day / SOLUTION LINK=LOGIT DIST=BINARY DDFM=BW;
      RANDOM INTERCEPT / TYPE=UN SUBJECT=patient_ID;
      ESTIMATE "Predicted Intercept at Day 0" intercept 1 day 0 / ILINK;
      ESTIMATE "Predicted Intercept at Day 1" intercept 1 day 1 / ILINK;
      ESTIMATE "Predicted Intercept at Day 2" intercept 1 day 2 / ILINK;
      ESTIMATE "Nonsense ILINK (un-logit) for day slope"  day 1 / ILINK;
RUN; TITLE3;

 * STATA Add Fixed Linear Slope for Day
melogit cam011 c.day,  || PATIENT_ID:  ,  ///
      covariance(unstructured) intpoints(7),
      estat ic, n(91),
      estimates store FixLin  // save for LRT
      lincom 1*_cons + 0*day  // predicted logit intercept at day 0
      lincom 1*_cons + 1*day  // predicted logit intercept at day 1
      lincom 1*_cons + 2*day  // predicted logit intercept at day 1
      nlcom 1/(1+(exp(-1*(1*_b[_cons] + 0*_b[day])))) // predicted prob at day 1
      nlcom 1/(1+(exp(-1*(1*_b[_cons] + 1*_b[day])))) // predicted prob at day 1
      nlcom 1/(1+(exp(-1*(1*_b[_cons] + 2*_b[day])))) // predicted prob at day 2
```

SAS Output:

Level-1: $\text{Logit}\left[p\left(y_{ti}=1\right)\right]=\beta_{0i}+\beta_{1i}\left(\text{Day}_{ti}\right)$

Level-2:

Intercept: $\beta_{0i}=\gamma_{00}+U_{0i}$

Linear Day: $\beta_{1i}=\gamma_{10}$

Fit Statistics

-2 Log Likelihood	625.16
AIC (smaller is better)	631.16
AICC (smaller is better)	631.21
BIC (smaller is better)	638.69

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Gradient
UN(1,1)	PATIENT_ID	1.5869	0.5137	0.000025

Note that the random intercept variance increases because day cannot reduce residual variance…

Solutions for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient
Intercept	0.9553	0.2418	90	3.95	0.0002	6.434E-6
day	-0.05670	0.04014	418	-1.41	0.1585	-0.00002

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean (In Probability)	Standard Error Mean (In Probability)
Predicted Intercept at Day 0	0.9553	0.2418	418	3.95	<.0001	0.7222	0.04852
Predicted Intercept at Day 1	0.8986	0.2182	418	4.12	<.0001	0.7107	0.04487
Predicted Intercept at Day 2	0.8419	0.2000	418	4.21	<.0001	0.6989	0.04208
Nonsense ILINK for day slope	-0.05670	0.04014	418	-1.41	0.1585	0.4858	0.01003

NOTE: You can "un-logit" (through the inverse link) model-predicted outcomes to get predicted probabilities. But you CANNOT "un-logit" slopes that represent a one-unit in logits (which is NOT the same as a one-unit change in probability). Here we see that the intercept distance between days (at −0.05670) is equal in logits, but this does not translate into equal distances in probability.

Model 3: Adding a Random Linear Slope for Days Since Hospital Admission (0=Day Hospitalized)

```
TITLE3 "SAS Add Random Linear Slope for Day";
PROC GLIMMIX DATA=Example5 NOCLPRINT NOITPRINT NAMELEN=100 GRADIENT METHOD=QUAD(QPOINTS=7);
       CLASS patient_ID;
       MODEL cam011 (DESCENDING) = day / SOLUTION LINK=LOGIT DIST=BINARY DDFM=BW;
       RANDOM INTERCEPT day / TYPE=UN SUBJECT=patient_ID;
       COVTEST "Need Random Linear Slope?"  . 0 0;
RUN; TITLE3;

 * STATA Add Fixed Linear Slope for Day
melogit cam011 c.day,  || PATIENT_ID: day ,  ///
       covariance(unstructured) intpoints(7),
       estimates store RandLin  // save for LRT
       lrtest RandLin FixLin    // test random linear
```

The . refers to the contents of the **G** matrix IN ORDER to be held to zero in estimating the comparison model.

SAS Output:

$$\text{Level-1: Logit}\left[p\left(y_{ti}=1\right)\right]=\beta_{0i}+\beta_{1i}\left(\text{Day}_{ti}\right)$$
Level-2:
$$\text{Intercept: } \beta_{0i}=\gamma_{00}+U_{0i}$$
$$\text{Linear Day: } \beta_{1i}=\gamma_{10}+U_{1i}$$

```
           Fit Statistics
-2 Log Likelihood             614.72
AIC  (smaller is better)      624.72
AICC (smaller is better)      624.84
BIC  (smaller is better)      637.27

            Covariance Parameter Estimates
Cov                                  Standard
Parm       Subject       Estimate       Error    Gradient
UN(1,1)    PATIENT_ID      1.6729      1.0415    -0.00022  random intercept variance
UN(2,1)    PATIENT_ID     -0.2037      0.2209    -0.00257  intercept-linear covariance
UN(2,2)    PATIENT_ID      0.1140     0.07322     -0.0072  random linear slope variance

                         Solutions for Fixed Effects
                         Standard
Effect        Estimate      Error       DF     t Value     Pr > |t|     Gradient
Intercept       0.9485     0.2693       90        3.52       0.0007      0.00028
day           -0.06101    0.07079      418       -0.86       0.3892     0.000834

                         Tests of Covariance Parameters
                             Based on the Likelihood
Label                          DF     -2 Log Like       ChiSq     Pr > ChiSq     Note
Need Random Linear Slope?       2          625.16       10.44         0.0033     MI
MI: P-value based on a mixture of chi-squares.  --: Standard test with unadjusted p-values.
```

COVTEST results: −2LL given is for a model without the random linear slope variance and the int-slope covariance.

Model 4: Adding a Fixed Quadratic Slope for Days Since Hospital Admission

$$\text{Level-1: Logit}\left[p\left(y_{ti}=1\right)\right]=\beta_{0i}+\beta_{1i}\left(\text{Day}_{ti}\right)+\beta_{2i}\left(\text{Day}_{ti}\right)^2$$
Level-2:
$$\text{Intercept: } \beta_{0i}=\gamma_{00}+U_{0i}$$
$$\text{Linear Day: } \beta_{1i}=\gamma_{10}+U_{1i}$$
$$\text{Quadratic Day: } \beta_{2i}=\gamma_{20}$$

```
TITLE3 "SAS Add Fixed Quadratic Slope for Day";
PROC GLIMMIX DATA=Example5 NOCLPRINT NOITPRINT NAMELEN=100 GRADIENT METHOD=QUAD(QPOINTS=7);
       CLASS patient_ID;
       MODEL cam011 (DESCENDING) = day day*day / SOLUTION LINK=LOGIT DIST=BINARY DDFM=BW;
       RANDOM INTERCEPT day / TYPE=UN SUBJECT=patient_ID;
RUN; TITLE3;
```

```
 * STATA Add Fixed Quadratic Slope for Day
melogit cam011 c.day c.daysq,  || PATIENT_ID: day ,  ///
       covariance(unstructured) intpoints(7),
       estat ic, n(91),
       estimates store FixQuad    // save for LRT
```

SAS Output:

```
           Fit Statistics
-2 Log Likelihood            614.58
AIC  (smaller is better)     626.58
AICC (smaller is better)     626.75
BIC  (smaller is better)     641.64

             Covariance Parameter Estimates
Cov                                  Standard
Parm       Subject        Estimate      Error    Gradient
UN(1,1)    PATIENT_ID       1.7064     1.0538    -0.00012  random intercept variance
UN(2,1)    PATIENT_ID      -0.2092     0.2219    0.000153  intercept-linear covariance
UN(2,2)    PATIENT_ID       0.1130    0.07273    0.002498  random linear slope variance

                        Solutions for Fixed Effects
                        Standard
Effect       Estimate      Error       DF    t Value    Pr > |t|    Gradient
Intercept      1.0263     0.3439       90       2.98      0.0037    0.000134
day           -0.1110     0.1522      417      -0.73      0.4664    -0.00162
day*day      0.005543    0.01492      417       0.37      0.7104     -0.0088
```

Model 5: Adding a Random Quadratic Slope for Days Since Hospital Admission

Level-1: $\text{Logit}\left[p\left(y_{ti}=1\right)\right]=\beta_{0i}+\beta_{1i}\left(\text{Day}_{ti}\right)+\beta_{2i}\left(\text{Day}_{ti}\right)^2$

Level-2:

Intercept: $\beta_{0i}=\gamma_{00}+U_{0i}$

Linear Day: $\beta_{1i}=\gamma_{10}+U_{1i}$

Quadratic Day: $\beta_{2i}=\gamma_{20}+U_{2i}$

```
TITLE3 "SAS Add Fixed Quadratic Slope for Day";
PROC GLIMMIX DATA=Example5 NOCLPRINT NOITPRINT NAMELEN=100 GRADIENT METHOD=QUAD(QPOINTS=7);
       CLASS patient_ID;
       MODEL cam011 (DESCENDING) = day day*day / SOLUTION LINK=LOGIT DIST=BINARY DDFM=BW;
       RANDOM INTERCEPT day day*day / TYPE=UN SUBJECT=patient_ID;
       COVTEST "Need Random Quadratic Slope?"  . . . 0 0 0;
RUN; TITLE3;

 * STATA Add Fixed Quadratic Slope for Day
melogit cam011 c.day c.daysq,  || PATIENT_ID: day daysq,  ///
       covariance(unstructured) intpoints(7),
       estat ic, n(91),
       estimates store RandQuad   // save for LRT
       lrtest RandQuad FixQuad    // test random quadratic
```

SAS Output:

```
           Fit Statistics
-2 Log Likelihood            609.09
AIC  (smaller is better)     627.09
AICC (smaller is better)     627.45
BIC  (smaller is better)     649.68
```

```
            Covariance Parameter Estimates
Cov                                       Standard
Parm        Subject        Estimate          Error     Gradient
UN(1,1)     PATIENT_ID       6.3279         3.9098     -0.00001
UN(2,1)     PATIENT_ID      -2.2355         1.5151     -0.00007
UN(2,2)     PATIENT_ID       1.0617         0.6515     -0.00009
UN(3,1)     PATIENT_ID       0.1656         0.1241     -0.00012
UN(3,2)     PATIENT_ID     -0.07685        0.05332     -0.00335
UN(3,3)     PATIENT_ID     0.005829       0.004559     -0.03605
                         Solutions for Fixed Effects
                         Standard
Effect         Estimate     Error      DF    t Value    Pr > |t|    Gradient
Intercept        1.3593    0.5098      90       2.67      0.0091    0.000013
day             -0.2586    0.2339     417      -1.11      0.2695    0.000013
day*day         0.02003   0.02231     417       0.90      0.3698    -0.00022

                           Tests of Covariance Parameters
                              Based on the Likelihood
Label                              DF    -2 Log Like      ChiSq    Pr > ChiSq     Note
Need Random Quadratic Slope?        3         633.88      24.80        <.0001     --
Need Random Linear Slope?           5         624.79      15.71        0.0077     --
Need Random Intercept?              6         671.49      62.40        <.0001     --
--: Standard test with unadjusted p-values.
```

SAS got this wrong! However, STATA got it closer to right: $-2\Delta LL(3) = 5.49$, $p = .139$.

Model 6: Adding Effects of MMSE (Level-2 Predictor) on Intercept and Linear Slope

$$\text{Level-1: Logit}\left[p\left(y_{ti}=1\right)\right]=\beta_{0i}+\beta_{1i}\left(\text{Day}_{ti}\right)$$

Level-2:

$$\text{Intercept:} \quad \beta_{0i}=\gamma_{00}+\gamma_{01}\left(\text{MMSE}_i-16\right)+U_{0i}$$

$$\text{Linear Day:} \quad \beta_{1i}=\gamma_{10}+\gamma_{11}\left(\text{MMSE}_i-16\right)+U_{1i}$$

```
TITLE3 "SAS Add Effects of MMSE on Intercept and Linear Slope";
PROC GLIMMIX DATA=Example5 NOCLPRINT NOITPRINT NAMELEN=100 GRADIENT METHOD=QUAD(QPOINTS=7);
       CLASS patient_ID;
       MODEL cam011 (DESCENDING) = day mmse16 day*mmse16 / SOLUTION LINK=LOGIT DIST=BINARY DDFM=BW;
       RANDOM INTERCEPT day / TYPE=UN SUBJECT=patient_ID;
RUN; TITLE3;

* STATA Add Effects of MMSE on Intercept and Linear Slope
xtmelogit cam011 c.day c.mmse16 c.mmse16#c.day, ||  PATIENT_ID: day ,   ///
       variance covariance(unstructured) intpoints(7),
       estat ic, n(91)
```

SAS Output:

```
           Fit Statistics
-2 Log Likelihood             583.67
AIC  (smaller is better)      597.67
AICC (smaller is better)      597.90
BIC  (smaller is better)      615.25

             Covariance Parameter Estimates
Cov                                       Standard
Parm        Subject        Estimate          Error     Gradient
UN(1,1)     PATIENT_ID       1.2946         0.8896     0.000134
UN(2,1)     PATIENT_ID      -0.2450         0.2061     0.000937
UN(2,2)     PATIENT_ID       0.1044        0.06443     0.000276
```

```
                          Solutions for Fixed Effects
                          Standard
Effect         Estimate      Error       DF    t Value    Pr > |t|     Gradient
Intercept        0.8444     0.2592       89       3.26      0.0016      0.00035
day             -0.1128    0.07094      417      -1.59      0.1127     0.000623
mmse16         -0.07268    0.03975       89      -1.83      0.0708     -0.00145
day*mmse16     -0.02022    0.01167      417      -1.73      0.0838     -0.00819
```

There is a *linear* relationship between day and the *logit* of at least some delirium.
There is a *nonlinear* relationship between day and the *probability* of at least some delirium.