Generalized Mixed Models for Ordinal Longitudinal Outcomes using PROC GLIMMIX

SAS Data Manipulation:

```
* Reading in all data;
DATA alldata; SET annk.annknewfinal; WHERE NMISS(age80, mmse16)=0;
      cam012=cam; LABEL cam012= "cam012: None, Some, Full Delirium"; run;
* Observed frequencies;
PROC FREQ DATA= alldata; TABLE cam012; run;
```

cam012: None, Some, Full Delirium

cam012	Frequency	Percent	Cumulative Frequency	Cumulative Percent
none 0	190	37.25	190	37.25
some 1	239	46.86	429	84.12
full 2	81	15.88	510	100.00

1a) Ordinal Single-Level Empty Means Model for None vs. Some vs. Full Delirium via GLIMMIX (Composite Model Form)—Note: it appears ILINK does not work correctly for CLOGIT models.

$$\text{Logit}(y_{ti} > 0) = \gamma_{001}$$
$$\text{Logit}(y_{ti} > 1) = \gamma_{002}$$

```
TITLE1 "GLIMMIX Ordinal Single-Level Empty Means Model";
TITLE2 "Cumulative Logit Link, Multinomial Distribution";
PROC GLIMMIX DATA=alldata METHOD=QUAD;
       CLASS patient_ID;
       MODEL cam012 (DESCENDING) = / SOLUTION LINK=CLOGIT DIST=MULT; run;
```

Response Profile

Ordered Value	cam012	Total Frequency
1	2	81
2	1	239
3	0	190

The GLIMMIX procedure is modeling the probabilities of levels of cam012 having lower Ordered Values in the Response Profile table.
This means we are predicting up, not down….

Convergence criterion (ABSGCONV=0.00001) satisfied.

Fit Statistics

-2 Log Likelihood	1035.58
AIC (smaller is better)	1039.58
AICC (smaller is better)	1039.60
BIC (smaller is better)	1048.05
CAIC (smaller is better)	1050.05
HQIC (smaller is better)	1042.90

Parameter Estimates

Effect	cam012: None, Prodromal, Full	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient
Intercept	2	-1.6670	0.1211	508	-13.76	<.0001	1.22E-14
Intercept	1	0.5213	0.09159	508	5.69	<.0001	-71E-15

Intercept2 = logit of probability of y > 1 = exp(−1.667) / (1+exp(−1.667)) = .1588

Intercept1 = logit of probability of y > 0 = exp(0.521) / (1+exp(0.521)) = .6274 (=.4686+.1588)

Probability(y=0) = 1 – intercept1 = 1 − .6274 = .3725
Probability(y=1) = intercept1 – intercept2 = .6274 − .1588 = .4686
Probability(y=2) = intercept2 − 0 = .1588 − 0 = .1588

1b) Ordinal Empty Means, Random Intercept Only Model (same random intercept both sub-models)

$$\text{Logit}(y_{ti} > 0) = \gamma_{001} + U_{0i}$$
$$\text{Logit}(y_{ti} > 1) = \gamma_{002} + U_{0i}$$

```
TITLE1 "GLIMMIX Ordinal Empty Means, Random Intercept Model";
TITLE2 "Cumulative Logit Link, Multinomial Response Distribution";
PROC GLIMMIX DATA=alldata NOCLPRINT NOITPRINT METHOD=QUAD(QPOINTS=15) GRADIENT;
       CLASS patient_ID;
       MODEL cam012 (DESCENDING) =  / SOLUTION LINK=CLOGIT DIST=MULT;
       RANDOM INTERCEPT / TYPE=UN SUBJECT=patient_ID;
              COVTEST "Need Random Intercept?"  0; * Test if G matrix (1,1)=0; RUN;
run;
```

```
          Fit Statistics
-2 Log Likelihood            978.26
AIC  (smaller is better)     984.26
AICC (smaller is better)     984.31
BIC  (smaller is better)     991.79
CAIC (smaller is better)     994.79
HQIC (smaller is better)     987.30

            Covariance Parameter Estimates
Cov                                  Standard
Parm       Subject       Estimate       Error    Gradient
UN(1,1)    PATIENT_ID      1.2547      0.3480    -0.00002


                            Solutions for Fixed Effects
            cam012:
            None,
            Prodromal,                Standard
Effect      Full          Estimate       Error       DF    t Value    Pr > |t|    Gradient
Intercept   2              -2.0423      0.1920       90     -10.63      <.0001    -0.00008
Intercept   1               0.6855      0.1640       90       4.18      <.0001    0.000062

                       Tests of Covariance Parameters
                          Based on the Likelihood
Label                       DF    -2 Log Like      ChiSq    Pr > ChiSq    Note

Need Random Intercept?       1        1035.58      57.32        <.0001    MI  This matches the -2ΔLL
MI: P-value based on a mixture of chi-squares.
```

The fixed effects for the intercepts are not the same as in the previous model. This is for two reasons:

(1) They are conditional on the random intercept (i.e., expected proportions for someone with $U_{0i} = 0$).
So for that kind of person, the probability of y > 1 = exp(−2.0411) / (1+exp(−2.0411)) = .115 (not .159),
and the probability of y > 0 = exp(0.6855) / (1+exp(0.06855)) = .665 (not .627).

(2) They are scaled differently because there is now more total variance in the outcome:
Total variance = $\text{Var}(U_{0i}) + \text{Var}(e_{ti}) = 12547 + \pi^2/3 = 1.2547 + 3.29 = 4.539$
= only 3.29 in previous single-level model

1c) Ordinal Mixed Model: Adding a Fixed Slope and a Random Slope for Days Since Hospital Admission (0=Day Hospitalized)

$$\text{Logit}(y_{ti} > 0) = \gamma_{001} + (\gamma_{10} + U_{1i})(\text{Day}_{ti}) + U_{0i}$$
$$\text{Logit}(y_{ti} > 1) = \gamma_{002} + (\gamma_{10} + U_{1i})(\text{Day}_{ti}) + U_{0i}$$

```
TITLE1 "GLIMMIX Ordinal Mixed Model -- Adding a Fixed and Random Slope for Day";
TITLE2 "Cumulative Logit Link, Multinomial Response Distribution";
PROC GLIMMIX DATA=alldata NOCLPRINT NOITPRINT METHOD=QUAD(QPOINTS=15) GRADIENT;
      CLASS patient_ID;
      MODEL cam012 (DESCENDING) = day / SOLUTION LINK=CLOGIT DIST=MULT;
      RANDOM INTERCEPT day / TYPE=UN SUBJECT=patient_ID;
      COVTEST "Need Random Slope?"  . 0 0; * Leave (1,1), test if (2,1) and (2,2) =0;
run;
```

Fit Statistics	
-2 Log Likelihood	966.27
AIC (smaller is better)	978.27
AICC (smaller is better)	978.43
BIC (smaller is better)	993.33
CAIC (smaller is better)	999.33
HQIC (smaller is better)	984.34

Fit statistics from Fixed Linear Day, Random Intercept Model:

Fit Statistics	
-2 Log Likelihood	973.70
AIC (smaller is better)	981.70
AICC (smaller is better)	981.78
BIC (smaller is better)	991.75
CAIC (smaller is better)	995.75
HQIC (smaller is better)	985.75

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Gradient
UN(1,1)	PATIENT_ID	1.9373	0.7856	-0.00005
UN(2,1)	PATIENT_ID	-0.1871	0.1484	-0.0004
UN(2,2)	PATIENT_ID	0.06592	0.04296	-0.00092

Solutions for Fixed Effects

Effect	cam012: None, Prodromal, Full	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient
Intercept	2	-1.8576	0.2582	90	-7.20	<.0001	-0.0002
Intercept	1	1.0682	0.2437	90	4.38	<.0001	0.000104
day		-0.09879	0.05501	84	-1.80	0.0761	-0.00088

Tests of Covariance Parameters Based on the Likelihood

Label	DF	-2 Log Like	ChiSq	Pr > ChiSq	Note
Need Random Slope?	2	973.70	7.44	0.0153	MI

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

This matches the $-2\Delta LL$

Interpret the fixed intercept 2:

Interpret the fixed intercept 1:

Interpret the fixed linear effect of day:

2a) Ordinal Mixed Model: Effects of MMSE on Intercept and Slope

$$\text{Logit}(y_{ti} > 0) = \gamma_{001} + (\gamma_{10} + U_{1i})(\text{Day}_{ti}) + \gamma_{01}(\text{MMSE}_i - 16) + \gamma_{11}(\text{MMSE}_i - 16)(\text{Day}_{ti}) + U_{0i}$$
$$\text{Logit}(y_{ti} > 1) = \gamma_{002} + (\gamma_{10} + U_{1i})(\text{Day}_{ti}) + \gamma_{01}(\text{MMSE}_i - 16) + \gamma_{11}(\text{MMSE}_i - 16)(\text{Day}_{ti}) + U_{0i}$$

```
TITLE1 "GLIMMIX Ordinal Mixed Model -- Effects of MMSE on Intercept and Slope";
TITLE2 "Cumulative Logit Link, Multinomial Response Distribution";
PROC GLIMMIX DATA=alldata NOCLPRINT NOITPRINT METHOD=QUAD(QPOINTS=15) GRADIENT;
       CLASS patient_ID;
       MODEL cam012 (DESCENDING) = day mmse16 day*mmse16 / SOLUTION LINK=CLOGIT DIST=MULT;
       RANDOM INTERCEPT day  / TYPE=UN SUBJECT=patient_ID;
       COVTEST "Still Need Random Intercept?"  0 0 0;
       COVTEST "Still Need Random Slope?"      . 0 0;  run;
```

```
          NOTE: GCONV convergence criterion satisfied.
FROM THE LOG:
NOTE: At least one element of the (projected) gradient is greater than 1e-3.
           Fit Statistics
-2 Log Likelihood             941.85
AIC  (smaller is better)      957.85
AICC (smaller is better)      958.14
BIC  (smaller is better)      977.94
CAIC (smaller is better)      985.94
HQIC (smaller is better)      965.96

            Covariance Parameter Estimates
Cov                                  Standard
Parm       Subject       Estimate       Error    Gradient
UN(1,1)    PATIENT_ID      1.5528      0.6970     -0.0012
UN(2,1)    PATIENT_ID     -0.1970      0.1418    -0.00588
UN(2,2)    PATIENT_ID     0.06289     0.04045    -0.00421

                              Solutions for Fixed Effects
             cam012:
             None,
             Prodromal,                Standard
Effect       Full          Estimate       Error       DF    t Value    Pr > |t|    Gradient
Intercept    2              -2.0024      0.2614       89      -7.66      <.0001    -0.00085
Intercept    1               0.9388      0.2427       89       3.87      0.0002     0.00126
day                         -0.1255     0.05860       83      -2.14      0.0352    0.007006
mmse16                     -0.08364     0.03233      333      -2.59      0.0101    -0.01616
day*mmse16                 -0.00822    0.008001      333      -1.03      0.3050    -0.05209

                         Tests of Covariance Parameters
                            Based on the Likelihood
Label                            DF    -2 Log Like      ChiSq    Pr > ChiSq    Note
Still Need Random Intercept?      3         986.88      45.03        <.0001    --
Still Need Random Slope?          2         949.46       7.61        0.0141    MI
MI: P-value based on a mixture of chi-squares.  --: Standard test with unadjusted p-values.
```

Interpret the fixed main effect of MMSE:

Interpret the fixed interaction of MMSE by Day:

This model makes what is referred as the "proportional odds" assumption. This means that although we have estimated separate intercepts for the two sub-models, we have constrained the fixed slopes for the predictor effects to be the same for each submodel (as well as the random effects). One alternative is the nominal model, which estimates separate intercepts, fixed effects, and random effects variances for each sub-model (specified as a baseline category vs. each of the other alternatives).

2b) Attempting the Nominal Model: Sub-models Prob(y=1 if 0 or 1, y=2 if 0 or 2)

$$\text{Logit}(y_{ti}>0)=\gamma_{001}+(\gamma_{101}+U_{1i1})(\text{Day}_{ti})+\gamma_{011}(\text{MMSE}_i-16)+\gamma_{111}(\text{MMSE}_i-16)(\text{Day}_{ti})+U_{0i1}$$
$$\text{Logit}(y_{ti}>1)=\gamma_{002}+(\gamma_{102}+U_{1i2})(\text{Day}_{ti})+\gamma_{012}(\text{MMSE}_i-16)+\gamma_{112}(\text{MMSE}_i-16)(\text{Day}_{ti})+U_{0i2}$$

```
TITLE1 "GLIMMIX Nominal Mixed Model -- Effects of MMSE on Intercept and Slope";
TITLE2 "Baseline Logit Link, Multinomial Response Distribution";
PROC GLIMMIX DATA=alldata NOCLPRINT NOITPRINT METHOD=QUAD(QPOINTS=15) GRADIENT;
       CLASS patient_ID cam012;
       MODEL cam012 (REF=FIRST) = day mmse16 day*mmse16 / SOLUTION LINK=GLOGIT DIST=MULT;
       RANDOM INTERCEPT day / TYPE=UN SUBJECT=patient_ID GROUP=cam012;
run;
```

```
The initial estimates did not yield a valid objective function.
```

Covariance Parameter Estimates

Cov Parm	Subject	Group	Estimate	Standard Error	Gradient
UN(1,1)	PATIENT_ID	cam012 1	1.8471	.	.
UN(2,1)	PATIENT_ID	cam012 1	-0.3055	.	.
UN(2,2)	PATIENT_ID	cam012 1	0.05900	.	.
UN(1,1)	PATIENT_ID	cam012 2	3.9296	.	.
UN(2,1)	PATIENT_ID	cam012 2	-0.4456	.	.
UN(2,2)	PATIENT_ID	cam012 2	0.07355	.	.

We have another alternative, in which the fixed effects of predictors are allowed to vary across submodels, but the random effects variances do not.

2c) Ordinal Mixed Model: NON-PROPORTIONAL Effects of MMSE on Intercept and Slope

$$\text{Logit}(y_{ti}>0)=\gamma_{001}+(\gamma_{101}+U_{1i})(\text{Day}_{ti})+\gamma_{011}(\text{MMSE}_i-16)+\gamma_{111}(\text{MMSE}_i-16)(\text{Day}_{ti})+U_{0i}$$
$$\text{Logit}(y_{ti}>1)=\gamma_{002}+(\gamma_{102}+U_{1i})(\text{Day}_{ti})+\gamma_{012}(\text{MMSE}_i-16)+\gamma_{112}(\text{MMSE}_i-16)(\text{Day}_{ti})+U_{0i}$$

```
TITLE1 "Ordinal Mixed Non-Proportional Model for MMSE";
PROC NLMIXED DATA=alldata METHOD=GAUSS TECH=QUANEW GCONV=1e-12;
       * Must list all parms to be estimated here with start values;
       * B01 and B02 = intercepts for each equation;
       * B's = fixed effects, now separate per equation;
       * V's = variance components in order of G matrix;
       PARMS  B01=.6 B02=-1.6
              B11day=0 B21mmse=0 B31mmseday=0
              B12day=0 B22mmse=0 B32mmseday=0
              V11=1 V21=-.2 V22=.05;
       * Linear predictive model;
       Y1 = B01 + B11day*day + U1*day + B21mmse*mmse16 + B31mmseday*mmse16*day + U0;
       Y2 = B02 + B12day*day + U1*day + B22mmse*mmse16 + B32mmseday*mmse16*day + U0;
       * Model for probability of response - writing it the shorter way;
            IF (cam=0) THEN P =  1 - (1/(1 + EXP(-Y1)));
       ELSE IF (cam=1) THEN P =      (1/(1 + EXP(-Y1))) - (1/(1 + EXP(-Y2)));
       ELSE IF (cam=2) THEN P =                           (1/(1 + EXP(-Y2)));
       LL = LOG(P);
       MODEL cam012 ~ GENERAL(LL);
       * Random intercept and linear slope;
       RANDOM U0 U1 ~ NORMAL([0,0],[V11,V21,V22]) SUBJECT=patient_ID;
RUN;
```

```
  NOTE: GCONV convergence criterion satisfied (and no error messages in the log!)

             Fit Statistics
-2 Log Likelihood                 932.9
AIC (smaller is better)           954.9
AICC (smaller is better)          955.4
BIC (smaller is better)           982.5
```

Fit statistics from Proportional version of same model:

```
           Fit Statistics
-2 Log Likelihood            941.85  -2ΔLL(3) = 9.05
AIC  (smaller is better)     957.85
AICC (smaller is better)     958.14
BIC  (smaller is better)     977.94
```

```
                                Parameter Estimates
                      Standard
Parameter    Estimate     Error    DF  t Value  Pr > |t|   Alpha     Lower     Upper  Gradient
B01            0.8466    0.2425    89     3.49    0.0008    0.05    0.3648    1.3284  1.426E-7
B02           -1.7277    0.3085    89    -5.60    <.0001    0.05   -2.3406   -1.1147   1.81E-8

B11day        -0.1059   0.06039    89    -1.75    0.0829    0.05   -0.2259   0.01409   -9.9E-7
B21mmse      -0.08176   0.03641    89    -2.25    0.0272    0.05   -0.1541  -0.00941  -8.25E-6
B31mmseday   -0.01454  0.009330    89    -1.56    0.1226    0.05  -0.03308  0.003996  -0.00004

B12day        -0.1515   0.07616    89    -1.99    0.0498    0.05   -0.3028  -0.00016  6.501E-7
B22mmse      -0.06759   0.03804    89    -1.78    0.0790    0.05   -0.1432  0.007991  1.199E-6
B32mmseday   -0.00254  0.009449    89    -0.27    0.7886    0.05  -0.02132   0.01623  6.219E-6

V11            1.2730    0.6449    89     1.97    0.0515    0.05  -0.00834    2.5543  1.418E-7
V21           -0.1614    0.1304    89    -1.24    0.2190    0.05   -0.4204   0.09765  1.014E-6
V22           0.06281   0.03859    89     1.63    0.1072    0.05  -0.01388    0.1395  9.457E-7
```

$-2\Delta LL(3) = 9.05$, $p < .05$, suggesting that each submodel needs its own set of fixed effects.
However, the interaction of MMSE*day is not significant for either submodel, and could be removed.

2d) Ordinal Mixed Non-Proportional Model: Removing MMSE*Day Interactions

$$\text{Logit}(y_{ti} > 0) = \gamma_{001} + (\gamma_{101} + U_{1i})(\text{Day}_{ti}) + \gamma_{011}(\text{MMSE}_i - 16) + U_{0i}$$
$$\text{Logit}(y_{ti} > 1) = \gamma_{002} + (\gamma_{102} + U_{1i})(\text{Day}_{ti}) + \gamma_{012}(\text{MMSE}_i - 16) + U_{0i}$$

```
TITLE1 "Ordinal Mixed Non-Proportional Model for MMSE - No MMSE*Day";
PROC NLMIXED DATA=alldata METHOD=GAUSS TECH=QUANEW GCONV=1e-12;
       * Must list all parms to be estimated here with start values;
       * B01 and B02 = intercepts for each equation;
       * B's = fixed effects;
       * V's = variance components in order of G matrix;
       PARMS  B01=.6 B02=-1.6 B11day=0 B21mmse=0 B12day=0 B22mmse=0
              V11=1 V21=-.2 V22=.05;
       * Linear predictive model - written as single-level equation;
       Y1 = B01 + U0 + B11day*day + B21mmse*mmse16 + U1*day;
       Y2 = B02 + U0 + B12day*day + B22mmse*mmse16 + U1*day;
       * Model for probability of response - writing it the shorter way;
            IF (cam=0) THEN P =  1 - (1/(1 + EXP(-Y1)));
       ELSE IF (cam=1) THEN P =      (1/(1 + EXP(-Y1))) - (1/(1 + EXP(-Y2)));
       ELSE IF (cam=2) THEN P =                           (1/(1 + EXP(-Y2)));
       LL = LOG(P);
       MODEL cam012 ~ GENERAL(LL);
       * Random intercept and linear slope;
       RANDOM U0 U1 ~ NORMAL([0,0],[V11,V21,V22]) SUBJECT=patient_ID;
RUN;
```

```
  NOTE: GCONV convergence criterion satisfied (and no error messages in the log!)

             Fit Statistics
-2 Log Likelihood                 936.2
AIC (smaller is better)           954.2
AICC (smaller is better)          954.6
BIC (smaller is better)           976.8
```

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
B01	0.8092	0.2408	89	3.36	0.0011	0.05	0.3307	1.2878	1.971E-6
B02	-1.7720	0.2927	89	-6.05	<.0001	0.05	-2.3536	-1.1903	-1.14E-6
B11day	-0.08586	0.05515	89	-1.56	0.1230	0.05	-0.1954	0.02372	0.000014
B21mmse	-0.1292	0.02455	89	-5.27	<.0001	0.05	-0.1780	-0.08047	-3.02E-6
B12day	-0.1314	0.06717	89	-1.96	0.0536	0.05	-0.2649	0.002058	-8.47E-6
B22mmse	-0.07113	0.02580	89	-2.76	0.0071	0.05	-0.1224	-0.01988	-0.00003
V11	1.3819	0.6610	89	2.09	0.0394	0.05	0.06859	2.6952	-1.46E-6
V21	-0.1654	0.1295	89	-1.28	0.2051	0.05	-0.4228	0.09203	-6.54E-6
V22	0.05569	0.03683	89	1.51	0.1340	0.05	-0.01749	0.1289	-0.00001