MLM 2007
Marginal vs RE models, Ordinal Responses (and other musings…)

Michael Griswold
Guest Lecture

Discussion Outline

- MLM review: Goals & Concepts
- Marginal & Random-Effect Models:
  - Logistic: PA & SS effects
  - Probit: PA & SS effects
  - Example: Crossover data (alcohol use)
- Ordinal Models
  - EDA
  - Extension of logistic regression (P.O. model)
  - Example: Schiz data (psychiatric drugs)
Multi-level Models: Review

Key Components of Multi-level Models

- Specification of predictor variables from multiple levels (fixed effects)
  - Variables to include
  - Key interactions
- Specification of correlation among responses from same clusters
  - Marginal (GEE)
  - Random (GLMM)
  - Transitional (Time-Series)
- Choices must be driven by scientific understanding, the research question and empirical evidence.
Digression on Statistical Models

- A statistical model is an approximation to reality
- There is not a “correct” model;
  - ( forget the holy grail )
- A model is a tool for asking a scientific question;
  - ( screw-driver vs. sludge-hammer )
- Useful models often combine the data with prior information to address the question of interest.
- Many models are better than one.

Multi-level Shmulti-level

- Multi-level analyses of social/behavioral phenomena: an important idea
- Multi-level models involve predictors from multiple-levels and their interactions
- They must account for associations among observations within clusters (levels) to make efficient and valid inferences.
Regression with Correlated Data

Must take account of correlation to:

- Obtain valid inferences
  - standard errors
  - confidence intervals
  - posteriors

- Make efficient inferences

Logistic Regression Example: Cross-over trial

<table>
<thead>
<tr>
<th>Group</th>
<th>(1,1)</th>
<th>(0,1)</th>
<th>(1,0)</th>
<th>(0,0)</th>
<th>Total</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>22</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>34</td>
<td>28</td>
<td>22</td>
</tr>
<tr>
<td>BA</td>
<td>18</td>
<td>4</td>
<td>2</td>
<td>9</td>
<td>33</td>
<td>20</td>
<td>22</td>
</tr>
</tbody>
</table>

- Response: 1-normal; 0- alcohol dependence
- Predictors: period ($x_1$); Placebo group ($x_2$)
- Two observations per person (cluster)
- Parameter of interest: log odds ratio of dependence: placebo VS treatment

Mean Model: $\log\{\text{odds(AD)}\} = \beta_0 + \beta_1 Period + \beta_2 Pl$
Marginal Models

- Focus is on the “mean model”: E(Y|X)
- Group comparisons are of main interest
  - Treatment vs non-treatment
  - Exposure vs non-exposure
  - Demographic comparisons
- Within-cluster associations are accounted for to correct standard errors, but are not of main interest.

Marginal Model Interpretations

- \( \log\{ \text{odds(AlcDep)} \} = \beta_0 + \beta_1 \text{Period} + \beta_2 \text{pl} \)
  - \( = 0.67 + (-0.30)\text{Period} + (0.57)\text{pl} \)
- TRT Effect: (placebo vs. trt)
  - \( \text{OR} = \exp(0.57) = 1.77, \quad 95\% \text{ CI (1.12, 2.80)} \)

\[ \text{Risk of Alcohol Dependence is almost twice as high on placebo, regardless of, (adjusting for), time period} \]

WHY?
Since: \( \log\{\text{odds(AlcDep|Period, pl)}\} = \beta_0 + \beta_1 \text{Period} + \beta_2 \)
And: \( \log\{\text{odds(AlcDep|Period, trt)}\} = \beta_0 + \beta_1 \text{Period} \)

\[ \Delta \log\text{-Odds} = \beta_2 \]
\[ \text{OR} = \exp(\beta_2) \]
**Random Effects Models**

- **Conditional on unobserved latent variables or “random effects”**
  - Responses (Alcohol use) within a person over time are usually related, but the association is not the same for everyone (heterogeneity)
  - Alcohol use within a family is related because family members share an unobserved “family effect”: common genes, diets, family culture and other unmeasured factors
  - Repeated observations within a neighborhood are correlated because neighbors share: common traditions, access to services, stress levels,…

---

**Random Effects Model Interpretations**

**WHY?**

Since:  \[ \log\{\text{odds}(\text{AlcDep}|\text{Period, pl, } b_i) \} = \beta_0 + \beta_1\text{Period} + \beta_2 + b_i \]

And:  \[ \log\{\text{odds}(\text{AlcDep}|\text{Period, trt, } b_i) \} = \beta_0 + \beta_1\text{Period} + b_i \]

\[ \Delta \log\text{-Odds} = \beta_2 \]

\[ \text{OR} = \exp(\beta_2) \]

- In order to make comparisons we must keep the subject-specific latent effect \((b_i)\) the same.
- In a Cross-Over trial we have outcome data for each subject on both placebo & treatment
- In other study designs we may not.
We have evidence on the conditional contrast in a cross-over trial. What about usual parallel RCTs?

**Marginal vs. Random Effects Models**

- For **linear models**, regression coefficients in random effects models and marginal models are identical:
  
  \[
  \text{average of linear function} = \text{linear function of average}
  \]

- For **non-linear models**, (logistic, log-linear,...) coefficients have different meanings/values, and address different questions
  
  - Marginal models -> *population-average* parameters
  
  - Random effects models -> *cluster-specific* parameters
Marginal -vs- Random Intercept Models; Cross-over Example

<table>
<thead>
<tr>
<th>Variable</th>
<th>Ordinary Logistic Regression</th>
<th>Marginal (GEE) Logistic Regression</th>
<th>Random-Effect Logistic Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.66 (0.32)</td>
<td>0.67 (0.29)</td>
<td>2.2 (1.0)</td>
</tr>
<tr>
<td>Period</td>
<td>-0.27 (0.38)</td>
<td>-0.30 (0.23)</td>
<td>-1.0 (0.84)</td>
</tr>
<tr>
<td>Treatment</td>
<td>0.56 (0.38)</td>
<td>0.57 (0.23)</td>
<td>1.8 (0.93)</td>
</tr>
<tr>
<td>Log OR (assoc.)</td>
<td>0.0 (0.81)</td>
<td>3.56 (0.81)</td>
<td>5.0 (2.3)</td>
</tr>
</tbody>
</table>

Comparison of Marginal and Random Effect Logistic Regressions

- Regression coefficients in the random effects model are roughly 3.3 times as large

  - Marginal: population odds (prevalence with/prevalence without) of AlcDep is \( \exp(0.57) = 1.8 \) greater for placebo than on active drug; population-average parameter

  - Random Effects: a person’s odds of AlcDep is \( \exp(1.8) = 6.0 \) times greater on placebo than on active drug; cluster-specific, here person-specific, parameter

Which model is better? They ask different questions.
Relationship between Marginal and RE models

\[ P(y_{ij} = 1 | x_1, x_2) = \int P(y_{ij} = 1 | x_1, x_2, \xi_i) \phi(\xi_i; 0, \hat{\tau}^2) d\xi_i \]

We can obtain marginal probabilities from the individual level probabilities by integrating out the random effects

Marginalized Multilevel Models!

Probit Regression Example: Cross-over trial

- Response: 1-normal; 0- alcohol dependence
- Predictors:
  - period \((x_1)\);
  - Placebo group \((x_2)\)
- Two observations per person (cluster)
- Parameter of interest: log odds ratio of dependence: treatment vs placebo

Mean Model: \( \Phi^{-1}\{Pr(AD=1)\} = \beta_0 + \beta_1 period + \beta_2 PI \)
Marginal -vs- Random Intercept Models; Cross-over Probit Example

<table>
<thead>
<tr>
<th>Variable</th>
<th>Ordinary Probit Regression</th>
<th>Marginal (GEE) Probit Regression</th>
<th>Random-Effect Probit Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.61 (0.38)</td>
<td>0.60 (0.29)</td>
<td>1.38 (0.65)</td>
</tr>
<tr>
<td>Period</td>
<td>-0.18 (0.23)</td>
<td>-0.19 (0.14)</td>
<td>-0.45 (0.35)</td>
</tr>
<tr>
<td>Treatment</td>
<td>0.29 (0.23)</td>
<td>0.34 (0.14)</td>
<td>0.79 (0.37)</td>
</tr>
<tr>
<td>Log tau (assoc.)</td>
<td>0.0</td>
<td>“nuisance”</td>
<td>0.67 (0.18)</td>
</tr>
</tbody>
</table>

Marginalized Probit Model

\[
P(y_{ij} = 1 \mid x) = \int P(y_{ij} = 1 \mid x, \zeta_i) \phi(\zeta_i; 0, \tau^2) d\zeta_i
\]

\[
= \int \Phi(x\beta + \zeta_i) \phi(\zeta_i; 0, \tau^2) d\zeta_i
\]

\[
= \Phi\left(\frac{x\beta}{\sqrt{1 + \tau^2}}\right)
\]

Closed Form Solution!
### Marginal -vs- Random Intercept Models; Cross-over Probit Example

<table>
<thead>
<tr>
<th>Variable</th>
<th>Ordinary Probit Regression</th>
<th>Marginal (GEE) Probit Regression</th>
<th>Random-Effect Probit Regression</th>
<th>MMM $\frac{\beta_{RE}}{\sqrt{1+\tau^2}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.61 (0.38)</td>
<td>0.60 (0.29)</td>
<td>1.38 (0.65)</td>
<td>0.63</td>
</tr>
<tr>
<td>Period</td>
<td>-0.18 (0.23)</td>
<td>-0.19 (0.14)</td>
<td>-0.45 (0.35)</td>
<td>-0.20</td>
</tr>
<tr>
<td>Treatment</td>
<td>0.29 (0.23)</td>
<td>0.34 (0.14)</td>
<td>0.79 (0.37)</td>
<td>0.35</td>
</tr>
<tr>
<td>Log tau (assoc.)</td>
<td>0.0 “nuisance”</td>
<td>0.67 (0.18)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Construct Contrasts of Interest

- **RE model:**
  - $\Phi^{-1}(Pr(AD=1)) = \beta_0 + \beta_1 \text{period} + \beta_2 \text{Pl} + \zeta_i$
  - with: $\zeta_i \sim N(0, \tau^2)$
  - $Pr(AD=1) = \Phi((\beta_0 + \beta_1 \text{period} + \beta_2 \text{Pl}) / \sqrt{1+\tau^2})$
- **Marginal RR(Pl vs trt, period 1)**
  - $= \Phi((\beta_0 + \beta_2) / \sqrt{1+\tau^2}) / \Phi(\beta_0 / \sqrt{1+\tau^2})$
- **Marginal OR, etc…**
Key Points

- “Multi-level” Models:
  - Have covariates from many levels and their interactions
  - Acknowledge correlation among observations from within a level (cluster)
- Assumptions about the latent variables determine the nature of the within cluster correlations
- Information can be borrowed across clusters (levels) to improve individual estimates
- Goal: Group Comparisons => Marginal Models
- Goal: Describe Heterogeneity => RE Models

Marginalized Multilevel Models

- Allows group comparisons
- Allows description of heterogeneity
- Allows associations to be non-nuisance
- Full Likelihood (RE) model => MAR
- Best parts of all worlds
Ordinal Responses

Binary outcome: \( \Phi^{-1}\{\Pr(Y=1)\} = \beta_0 + X\beta \)

(At Centered X)
Latent Response (probit) form

Ordinal outcome:  \( \Phi^{-1}\{\Pr(Y>s)\} = \alpha_s + X\beta \)

Cumulative Response Models

- Logistic regression: 2-categories (0/1)
  - \( \log\{ \Pr(Y=1) / [1-\Pr(Y=1)] \} = \beta_0 + X\beta \)
  - \( \log\{ \Pr(Y=1) / \Pr(Y=0) \} = \beta_0 + X\beta \)
  - \( \log\{ \Pr(Y>0) / \Pr(Y<0) \} = \beta_0 + X\beta \)

- P.O. regression: S-categories (1,2,...,S)
  - \( \log\{ \Pr(Y>1) / \Pr(Y\leq1) \} = \alpha_1 + X\beta \)
  - \( \log\{ \Pr(Y>2) / \Pr(Y\leq2) \} = \alpha_2 + X\beta \)
  - \( \log\{ \Pr(Y>s) / \Pr(Y\leq s) \} = \alpha_s + X\beta \)
  - \( \log\{ \Pr(Y>s) / [1-\Pr(Y>s)] \} = \alpha_s + X\beta \)

- Note: Gllamm uses \(-k_s\) for \(\alpha_s\)
Ordered Responses

- **Probit:** \( \Phi^{-1}\{\Pr(Y>s)\} = \alpha_s + X\beta \)

- **PO:** \( \log\{\text{odds}(Y>s)\} = \alpha_s + X\beta \)

- \( s = 1..(S-1) \) & check manuals for \(-\alpha_s, -X\beta\)

- Interpretations: \( \beta \) represents the assoc of a 1-unit increase in \( X \) with a change in logodds of being in ANY cumulative cat.

- Ex: 3-cat PO: \( \log\{\text{odds}(Y>s)\} = \alpha_s + X\beta, \ s=1,2 \)
  - \( \log\text{odds}(Y>1) = \alpha_1 + X\beta \)
  - \( \log\text{odds}(Y>2) = \alpha_2 + X\beta \)

Schiz Data:

Schizophrenia Collaborative Study (NIMH)

- Antipsychotic Drugs & Schiz. Severity

- 437 patients
  - Placebo (0) & treatment (1)
  - Trt = (Chlorpromazine, Fluphenazine, or Thioridazine)

- 7 potential visits for each patient (0..6)

- Outcome: IMPS item 79
  - Inpatient Multidimensional Psychiatric Scale
  - 1=Normal, 2=mildly ill, 3=markedly ill, 4=severely

- Q1) How well does trt work vs Placebo?
- Q2) How variable are patients’ responses
Schiz Data cont: Data Patterns

<table>
<thead>
<tr>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
<th>Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>308</td>
<td>70.48</td>
<td>70.48</td>
<td>11.1..1</td>
</tr>
<tr>
<td>41</td>
<td>9.38</td>
<td>79.86</td>
<td>11.1...</td>
</tr>
<tr>
<td>37</td>
<td>8.47</td>
<td>88.33</td>
<td>11.....</td>
</tr>
<tr>
<td>8</td>
<td>1.83</td>
<td>90.16</td>
<td>11....1</td>
</tr>
<tr>
<td>8</td>
<td>1.83</td>
<td>91.99</td>
<td>111....</td>
</tr>
<tr>
<td>6</td>
<td>1.37</td>
<td>93.36</td>
<td>11.1.1.</td>
</tr>
<tr>
<td>5</td>
<td>1.14</td>
<td>94.51</td>
<td>1..1..1</td>
</tr>
<tr>
<td>5</td>
<td>1.14</td>
<td>95.65</td>
<td>11.11..</td>
</tr>
<tr>
<td>3</td>
<td>0.69</td>
<td>96.34</td>
<td>.1.1..1</td>
</tr>
<tr>
<td>16</td>
<td>3.66</td>
<td>100.0</td>
<td>(other patterns)</td>
</tr>
<tr>
<td>437</td>
<td>100.00</td>
<td>XXXXXX</td>
<td></td>
</tr>
</tbody>
</table>

Cumulative Probabilities over Weeks

Graphs by treatment
Marginal Cumulative Probabilities
Ordinary P.O. model: stata

ologit impso weeksqrt treatment interact, or robust cluster(id)

Ordered logistic regression  Number of obs =    1603
Wald chi2(3)  =   440.17
Prob > chi2 =    0.0000
Log pseudolikelihood = -1878.0969  Pseudo R2 =    0.1177

(Std. Err. adjusted for 437 clusters in id)

|                | Robust     | z   | P>|z|   | [95% Conf. Interval] |
|----------------|------------|-----|-------|---------------------|
|                |            |     |       |                     |
| impso          |            |     |       |                     |
| weeksqrt       | .5847056   | .0591797 | -5.30 | 0.000       | .4794958    .7130004 |
| treatment      | .9993959   | .2042595 | -0.00 | 0.998       | .6695244    1.491793  |
| interact       | .4719089   | .0568135 | -6.24 | 0.000       | .3727189    .5974961  |

| /cut1          | -3.807279  | .1956796 |       | -4.190804    -3.423754 |
| /cut2          | -1.760167  | .1811041 |       | -2.115125    -1.40521  |
| /cut3          | -4.221112  | .1795596 |       | -7.740415    -0.701808 |

Cumulative Probits over $\sqrt{\text{weeks}}$

Marginal Cumulative Probits
Ordinary PO interpretations

Model: \( \log\{\text{odds}(Y>s)\} = \alpha_s + \beta_1swk + \beta_2trt + \beta_3swk*trt \)

- \( \log\{\text{odds}(Y>1 | wk=0, trt)\} = \alpha_1 + \beta_2 \)
- \( \log\{\text{odds}(Y>1 | wk=0, Pl)\} = \alpha_1 \)
  \[ \exp(\beta_2) = 1.0 \]

- \( \log\{\text{odds}(Y>1 | wk=0, trt)\} = \alpha_1 + \beta_2 \)
- \( \log\{\text{odds}(Y>1 | wk=0, Pl)\} = \alpha_1 \)
  \[ \exp(\beta_2) = 1.0 \]

- Effects are the same across cumulative cats
- No effect at baseline

Ordinary PO interpretations

Model: \( \log\{\text{odds}(Y>s)\} = \alpha_s + \beta_1swk + \beta_2trt + \beta_3swk*trt \)

- \( \log\{\text{odds}(Y>1 | wk=1, trt)\} = \alpha_1 + \beta_1 + \beta_2 + \beta_3 \)
- \( \log\{\text{odds}(Y>1 | wk=1, Pl)\} = \alpha_1 + \beta_1 \)
  \[ \exp(\beta_2+\beta_3) = 0.28 \]

- \( \log\{\text{odds}(Y>2 | wk=1, trt)\} = \alpha_1 + \beta_1 + \beta_2 + \beta_3 \)
- \( \log\{\text{odds}(Y>2 | wk=1, Pl)\} = \alpha_1 + \beta_1 \)
  \[ \exp(\beta_2+\beta_3) = 0.28 \]

- Effects are the same across cumulative cats
- 72% Reduction in “risk” (odds) at wk1, trt vs pl
Observed & Predicted Probs

Graphs by treatment

Diagnostic Check looks good

Ordered Responses w/ Ran Ints

- Probit: $\Phi^{-1}\{\Pr(Y>s)\} = \alpha_s + X\beta + u_i$
- PO: $\log\{\text{odds}(Y>s)\} = \alpha_s + X\beta + u_i$
- $s = 1..(S-1)$ & check manuals for $-\alpha_s, -X\beta$
- Interpretations: $\beta$ represents the assoc of a 1-unit increase in $X$ with a change in logodds of being in ANY cumulative cat. for a single patient.
- Ex: 3-cat PO: $\log\{\text{odds}(Y>s)\} = \alpha_s + X\beta + u_i$, $s=1,2$
  - $\log\text{odds}(Y>1) = \alpha_1 + X\beta + u_i$
  - $\log\text{odds}(Y>2) = \alpha_2 + X\beta + u_i$  

Strong Assumption
Ran Int P.O. model: gllamm

gllamm impso weeksqrt treatment interact,
   i(id) link(ologit) adapt eform

------------------------------------------------------------------------------
|         exp(b)   Std. Err.      z    P>|z|     [95% Conf. Interval] |
|----------|----------------|------------|--------------|-----------------------------|
| impso    |                |             |              |                             |
| weeksqrt | 0.4649525      | 0.0608031  | -5.86        | 0.000                       | 0.3598277    0.6007899 |
| treatment| 0.9439404      | 0.2962807  | -0.18        | 0.854                       | 0.5102375    1.746291  |
| interact | 0.2993646      | 0.0457031  | -7.90        | 0.000                       | 0.2219474    0.4037855 |
------------------------------------------------------------------------------
|_cut11    | -5.858453      | 0.331792   | -17.66       | 0.000                       | -6.508753    -5.208153 |
------------------------------------------------------------------------------
|_cut12    | -2.825669      | 0.2900513  | -9.74        | 0.000                       | -3.394159    -2.257179 |
------------------------------------------------------------------------------
|_cut13    | -0.7077072     | 0.2750904  | -2.57        | 0.010                       | -1.246875    -0.1685399|
------------------------------------------------------------------------------

Variances and covariances of random effects

***level 2 (id)
var(1): 3.7733416 (.46496878)

------------------------------------------------------------------------------

Ran. Int. (SS) PO interpretations

Model: log\{odds(Y>s)} = \alpha_s + \beta_1 swk + \beta_2 trt + \beta_3 swk*trt + u_i

- log\{odds(Y>1 | wk=0, trt)} = \alpha_1 + \beta_2 + u_i
- log\{odds(Y>1 | wk=0, Pl)} = \alpha_1 + u_i

\[ \exp(\beta_2) = 0.94 \]

- log\{odds(Y>2 | wk=0, trt)} = \alpha_2 + \beta_2
- log\{odds(Y>2 | wk=0, Pl)} = \alpha_2

\[ \exp(\beta_2) = 0.94 \]

- At baseline, no effect comparing a single patient on trt, to that same patient off trt??
Ran. Int. (SS) PO interpretations

Model: \( \log\{\text{odds}(Y > s)\} = \alpha_s + \beta_1 \text{swk} + \beta_2 \text{trt} + \beta_3 \text{swk*trt} + u_i \)

- \( \log\{\text{odds}(Y > 1 | \text{wk} = 1, \text{trt})\} = \alpha_1 + \beta_1 + \beta_2 + \beta_3 + u_i \)
- \( \log\{\text{odds}(Y > 1 | \text{wk} = 1, \text{Pl})\} = \alpha_1 + \beta_1 + u_i \)

\( \exp(\beta_2 + \beta_3) = 0.14 \)

- Similar at week 1, etc. The SS trt effect compares a single patient on trt, to that same patient off trt but we have not observed any actual data on this effect. This is a “causal extrapolation”

How Heterogeneous is the data?

Model: \( \log\{\text{odds}(Y > s)\} = \alpha_s + \beta_1 \text{swk} + \beta_2 \text{trt} + \beta_3 \text{swk*trt} + u_i \)

- If a patient is on trt (or off), how variable is their specific outcome trajectory?
- Estimate of Ran Int variance: 3.77 (0.46)
- Huge!
- Can we visualize?
- Sure, use Empirical Bayes estimates of \( u_i \)
Patient-Specific Trajectories

Graphs by newid

Placebo

Patient-Specific Trajectories

Graphs by newid

Treatment
Can we Marginalize the PO model?

- Of course, the marginalized version integrates the random effects out over their assumed distribution
- no more causal extrapolation
- Currently not implemented in Stata, but see “A User Friendly Guide to Link-Probit Models” – Caffo, Griswold; TAS 2006
- We can use Gllamm’s post-estimation prediction to compute the marginal probabilities for visualization however…

Marginal Cum. Prob. Trajectories

Cumulative Probabilities
Marginal Prob. Trajectories

Graphs by treatment

Ordinal Category Probabilities

Relaxing the PO assumption

- **PO:** \( \log\{\text{odds}(Y>s)\} = \alpha_s + X\beta \)
- **Non-PO:** \( \log\{\text{odds}(Y>s)\} = \alpha_s + X\beta_s \)
- \( s = 1..(S-1) \) & check manuals for \(-\alpha_s, -X\beta_s\)
- Interpretations: \( \beta_s \) represents the assoc of a 1-unit increase in \( X \) with a change in logodds of being in cumulative cat. \( \text{“}s\text{”} \)
- **Ex: 3-cat PO:** \( \log\{\text{odds}(Y>s)\} = \alpha_s + X\beta_s, \ s=1,2 \)
  - \( \log\text{odds}(Y>1) = \alpha_1 + X\beta_{11} \)
  - \( \log\text{odds}(Y>2) = \alpha_2 + X\beta_{21} \) Relaxed Assumption
non-P.O. model: gologit

```
gologit impso weeksqrt treatment interact, cluster(id) robust
(Std. Err. adjusted for 437 clusters in id)

------------------------------------------------------------------------------
|               Robust          
| impos | Odds Ratio | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|--------|------------|-----------|------|-----|-------------------|
| mleq1  |            |           |      |     |                   |
| weeksqrt | 0.2041182  | 0.0937231 | -3.46| 0.001| 0.0829934 - 0.5020189 |
| treatment | 0.1793261  | 0.1833936 | -1.68| 0.093| 0.0241621 - 1.330921 |
| interact | 1.054759   | 0.5020053 | 0.11| 0.911| 0.4149823 - 2.680876 |

| mleq2  |            |           |      |     |                   |
| weeksqrt | 0.4847473  | 0.0609446 | -5.76| 0.000| 0.3788772 - 0.6202008 |
| treatment | 0.7877922  | 0.2234486 | -0.84| 0.400| 0.4518327 - 1.373554 |
| interact | 0.5892814  | 0.0865812 | -3.60| 0.000| 0.4418333 - 0.7859357 |

| mleq3  |            |           |      |     |                   |
| weeksqrt | 0.66977    | 0.069234  | -3.88| 0.000| 0.5469368 - 0.8201896 |
| treatment | 1.061012   | 0.2399695 | 0.26| 0.793| 0.6810892 - 1.652863 |
| interact | 0.441081   | 0.0588683 | -6.13| 0.000| 0.3395579 - 0.7529581 |

_cons1 | 5.986731   | 0.9860904 | 6.07| 0.000| 4.05403 - 7.919433  |
_cons2 | 1.996487   | 0.2504722 | 7.97| 0.000| 1.505571 - 2.487404  |
_cons3 | 3.0472     | 0.1993504 | 1.53| 0.126| -0.0859997 - 0.6954397 |
------------------------------------------------------------------------------
```

Compare w/ P.O. model: ologit

```
ologit impso weeksqrt treatment interact, or robust cluster(id)

Ordered logistic regression
Number of obs = 1603
Wald chi2(3)  = 440.17
Prob > chi2 = 0.0000
Log pseudolikelihood = -1878.0969 Pseudo R2 = 0.1177

<table>
<thead>
<tr>
<th></th>
<th>Robust</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>impos</td>
<td>Odds Ratio</td>
<td>Std. Err.</td>
<td>z</td>
<td>P&gt;</td>
<td>z</td>
</tr>
<tr>
<td>weeksqrt</td>
<td>0.5847056</td>
<td>0.0591797</td>
<td>-5.30</td>
<td>0.000</td>
<td>0.4794958 - 0.7130004</td>
</tr>
<tr>
<td>treatment</td>
<td>0.9993959</td>
<td>0.2042595</td>
<td>-0.00</td>
<td>0.998</td>
<td>0.6695244 - 1.491793</td>
</tr>
<tr>
<td>interact</td>
<td>0.4719089</td>
<td>0.0568135</td>
<td>-6.24</td>
<td>0.000</td>
<td>0.3727189 - 0.5974961</td>
</tr>
<tr>
<td>/cut1</td>
<td>-3.807279</td>
<td>0.1956796</td>
<td>-4.190804 - 3.423754</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/cut2</td>
<td>-1.760167</td>
<td>0.1811041</td>
<td>-2.115125 - 1.40521</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/cut3</td>
<td>-4.221112</td>
<td>0.1795596</td>
<td>-7.740415 - 0.0701808</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
Relaxing the PO assumption

- **PO:** \( \log\{\text{odds}(Y>s)\} = \alpha_s + X\beta + u_i \)
- **Non-PO:** \( \log\{\text{odds}(Y>s)\} = \alpha_s + X\beta_s + u_i \)
- \( s = 1..(S-1) \) & check manuals for \(-\alpha_s, -X\beta_s\)
- Interpretations: \( \beta_s \) represents the assoc of a 1-unit increase in \( X \) with a change in logodds of being in cumulative cat. “\( s \)” for a single patient.
- **Ex:** 3-cat PO: \( \log\{\text{odds}(Y>s)\} = \alpha_s + X\beta + u_i \), \( s=1,2 \)
  - \( \log\text{odds}(Y>1) = \alpha_1 + X\beta_1 + u_i \)
  - \( \log\text{odds}(Y>2) = \alpha_2 + X\beta_2 + u_i \)  
  
Ran. Int. (SS) non-PO interpretations

\[ \log\{\text{odds}(Y>s)\} = \alpha_s + \beta_{s1}\text{swk} + \beta_{s2}\text{trt} + \beta_{s3}\text{swk*trt} + u_i \]

- \( \log\text{odds}(Y>1 \mid \text{wk}=0, \text{trt}) = \alpha_1 + \beta_{12} + u_i \)
- \( \log\text{odds}(Y>1 \mid \text{wk}=0, \text{Pl}) = \alpha_1 + u_i \)  
  \[ \exp(\beta_{12}) = ?? \]

- \( \log\text{odds}(Y>2 \mid \text{wk}=0, \text{trt}) = \alpha_2 + \beta_{22} \)
- \( \log\text{odds}(Y>2 \mid \text{wk}=0, \text{Pl}) = \alpha_2 \)  
  \[ \exp(\beta_{22}) = ?? \]

- Gllamm still running…
Schiz Summary

- Under a common trt effect, general 72% decrease in cumulative odds risk per unit time (sqrt week).
- Patient responses are highly variable, so the marginal responses may not fit an individual’s response well.
- Could model this with MMM (probit) to handle both estimation aspects
- Potentially less change over time in lower categories
- Potentially stronger trt effects in upper categories

Ordinal MLM notes

- PO models are basically logistic regressions
  - popular
  - strong parallel regression assumption
  - Can be relaxed
- Mixed PO have SS, not PA effects (from logit)
- Other models:
  - Ordinal Probit
  - Continuation ratio model
  - Multinomial logit model
- Additional REs (random slopes, etc.)
From Caffo & Griswold TAS 2006:
Ordinal MMM with 2 REs

Contour plot of the fitted bivariate distribution
along with cell counts and fitted cell counts

Overall Summary: MLMs

- Powerful tools / dangerous black boxes
- “Buyer Beware”
  - Model Assumptions: both fixed AND random ($u_i \sim N(0, \tau^2)$)
  - Identifiability
  - Model Fit: Marginalize & Check whenever possible
  - Report Heterogeneity as well (& meaning)
  - MLMs require even more due-diligence than usual

- Marginal Models (~GEE)
  - Nice PA interpretations, more robust

- RE models (~GLMM)
  - Nice MAR, flexible assoc, full likelihood

- MMM: best of both worlds