Mixed-effects Regression Models
for Longitudinal Dichotomous Data

Chapter 9
Logistic Regression - model that relates explanatory variables (i.e., covariates) to a dichotomous dependent variable

Mixed-effects Logistic Regression - model that relates covariates to a dichotomous dependent variable, where observations are nested

- Longitudinal: repeated observations within subjects
- Clustered: subjects within clusters

models can also be recast as probit regression models
Logistic Regression Model with dichotomous $x$

<table>
<thead>
<tr>
<th>group</th>
<th>$x$</th>
<th>$Y = \text{response}$</th>
<th>prob</th>
<th>odds</th>
<th>logit</th>
</tr>
</thead>
<tbody>
<tr>
<td>control</td>
<td>0</td>
<td>60 30</td>
<td>1/3</td>
<td>1/2</td>
<td>-.693</td>
</tr>
<tr>
<td>treatment</td>
<td>1</td>
<td>30 60</td>
<td>2/3</td>
<td>2</td>
<td>.693</td>
</tr>
</tbody>
</table>

\[
\log \left[ \frac{Pr(Y_i = 1)}{1 - Pr(Y_i = 1)} \right] = \beta_0 + \beta_1 x_i
\]

\[
\exp \beta_0 = \text{odds of response for } x = 0 \quad (30/60 = 1/2)
\]
\[
\hat{\beta}_0 = \log(1/2) = -.693
\]

\[
\exp(\beta_0 + \beta_1) = \text{odds of response for } x = 1 \quad (60/30 = 2)
\]
\[
\hat{\beta}_0 + \hat{\beta}_1 = \log(2) = .693
\]
\[
\hat{\beta}_1 = .693 + .693 = 1.386
\]
odds ratio = ratio of odds per unit change in $x$

$$= \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1)}{\exp(\hat{\beta}_0)}$$

$$= \exp(\hat{\beta}_1)$$

$$= \exp(1.386) = 4$$
Model is not linear in terms of the probabilities

\[
Pr(Y_i = 1) = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 x_i)]} = \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}
\]
Model is linear in terms of the logits

\[
\log \left[ \frac{Pr(Y_i = 1)}{1 - Pr(Y_i = 1)} \right] = \beta_0 + \beta_1 x_i
\]
Logistic Regression Model with continuous $x$

<table>
<thead>
<tr>
<th>age</th>
<th>$x$</th>
<th>$Y = \text{response}$</th>
<th>prob</th>
<th>odds</th>
<th>logit</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-29</td>
<td>0</td>
<td>60 30</td>
<td>1/3</td>
<td>1/2</td>
<td>-.693</td>
</tr>
<tr>
<td>30-39</td>
<td>1</td>
<td>30 60</td>
<td>2/3</td>
<td>2</td>
<td>.693</td>
</tr>
<tr>
<td>40-49</td>
<td>2</td>
<td>10 80</td>
<td>8/9</td>
<td>8</td>
<td>2.079</td>
</tr>
</tbody>
</table>

$$\log \left[ \frac{Pr(Y_i = 1)}{1 - Pr(Y_i = 1)} \right] = \beta_0 + \beta_1 x_i$$

$$\hat{\beta}_0 = -.693$$

$$\hat{\beta}_1 = \text{change in log odds w/ unit change in } x$$

$$= 1.386$$
odds ratio = ratio of odds per unit change in $x$

= $\exp(\hat{\beta}_1)$

= $\exp(1.386) = 4$
\[ Pr(Y_i = 1) = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 x_i)]} = \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)} \]
Logistic regression in terms of logit

\[ \log \left( \frac{Pr(Y_i = 1)}{1 - Pr(Y_i = 1)} \right) = \beta_0 + \beta_1 x_i \]
ML estimation

\[ \Pr(Y_i) = \Psi_i^{Y_i}[1 - \Psi_i]^{1-Y_i} \quad \text{for } Y_i = 0 \text{ or } 1, \]

where \( \Psi_i = \Psi_i(x'_i/\beta) = \frac{1}{1+\exp(-x'_i/\beta)} \)

The likelihood function for a sample of \( N \) independent observations can be written as the product over the \( N \) individuals, \( i.e., \)

\[ L = \prod_{i=1}^{N} \Psi_i^{Y_i}[1 - \Psi_i]^{1-Y_i} \]

Thus the log-likelihood function becomes

\[ \log L = \sum_{i=1}^{N} [Y_i \log \Psi_i + (1 - Y_i) \log(1 - \Psi_i)] \]
Differentiating the log likelihood function with respect to $\beta$ yields the first derivatives for the maximum likelihood (ML) solution:

$$ \frac{\partial \log L}{\partial \beta} = \sum_i (Y_i - \Psi_i) x_i $$

This result is due to the fact that for the logistic distribution $\delta \Psi(\cdot) = \Psi(\cdot)(1 - \Psi(\cdot))$. Similarly, the second derivatives are obtained as

$$ \frac{\partial^2 \log L}{\partial \beta \partial \beta'} = -\sum_i \Psi_i (1 - \Psi_i) x_i x_i' $$

In the solution via Newton-Raphson, provisional estimates for the vector of parameters $\beta$, on iteration $i$ are improved by

$$ \beta_{i+1} = \beta_i - \left[ \frac{\partial^2 \log L}{\partial \beta \partial \beta'} \right]^{-1} \frac{\partial \log L}{\partial \beta_i} $$

until convergence
Random-intercept Logistic Regression Model

Consider the model with $p$ covariates for the dichotomous response $Y_{ij}$ of subject $i$ ($i = 1, \ldots, N$) at timepoint $j$ ($j = 1, \ldots, n_i$):

$$
\log \left[ \frac{Pr(Y_{ij} = 1)}{1 - Pr(Y_{ij} = 1)} \right] = \mathbf{x}_{ij}'\mathbf{\beta} + \upsilon_i
$$

$Y_{ij} = \text{dichotomous response of subject } i \text{ at timepoint } j$

$\mathbf{x}_{ij} = (p + 1) \times 1 \text{ vector of covariates}$

$\mathbf{\beta} = (p + 1) \times 1 \text{ vector of regression coefficients}$

$\upsilon_i = \text{random subject effects distributed } \mathcal{NID}(0, \sigma_{\upsilon}^2)$
Dichotomous Response and Threshold Concept

Continuous $y_{ij}$ - an unobservable latent variable - related to dichotomous response $Y_{ij}$ via “threshold concept”

- threshold value $\gamma$ on $y$ continuum

Response occurs $Y_{ij} = 1$ if $\gamma < y_{ij}$
otherwise, a response does not occur ($Y_{ij} = 0$)
The Threshold Concept in Practice

“How was your day?”
(What is your level of satisfaction today?)

- Satisfaction may be continuous, but we usually emit a dichotomous response:

  Great Day!

  a day ...
Model for Latent Continuous Responses

\[ y_{ij} = x'_{ij}\beta + v_i + \varepsilon_{ij} \]

- \( \varepsilon_{ij} \sim \text{std normal (mean 0, variance 1): probit regression} \)
- \( \varepsilon_{ij} \sim \text{std logistic (mean 0, variance } \pi^2/3): \text{logistic regression} \)

Underlying latent variable
- useful way of thinking of the problem
- not an essential assumption of the model
- used for intra-class correlation

\[ ICC = \frac{\sigma_v^2}{\sigma_v^2 + 1} \quad \text{for probit (equals tetrachoric if } n = 2) \]

\[ = \frac{\sigma_v^2}{\sigma_v^2 + \pi^2/3} \quad \text{for logistic} \]
Scaling of regression coefficients

*Fixed-effects or marginal model* - $\beta$ estimates from logistic are larger in absolute value than from probit by

$$\approx \sqrt{\frac{\pi^2/3}{1}} = \sqrt{\frac{\text{std logistic variance}}{\text{std normal variance}}} = 1.8$$


*Random-effects model* - $\beta$ estimates from random-effects model are larger in abs. value than fixed-effects or marginal model by

$$\approx \sqrt{d} = \sqrt{\frac{\sigma^2_y + \sigma^2}{\sigma^2}} = \sqrt{\frac{\text{RE variance}}{\text{FE variance}}}$$

- $d =$ design effect in sampling literature
- Zeger et. al. (1988) $\sigma^2 = (15/16)^2 \pi^2/3$ for logistic
Random-Intercept Model  *Within-Subjects / Between-Subjects models*

**Within-subjects model - level 1**  \((j = 1, \ldots, n_i)\)

\[
\begin{align*}
\text{observed response} & \\
\log \left[ \frac{Pr(Y_{ij} = 1)}{1 - Pr(Y_{ij} = 1)} \right] & = b_{0i} + b_{1i} Time_{ij}
\end{align*}
\]

\[
\begin{align*}
\text{latent response} & \\
y_{ij} & = b_{0i} + b_{1i} Time_{ij} + \varepsilon_{ij}
\end{align*}
\]

**Between-subjects model - level 2**  \((i = 1, \ldots, N)\)

\[
\begin{align*}
b_{0i} & = \beta_0 + \beta_2 Grp_i + \nu_{0i} \\
b_{1i} & = \beta_1 + \beta_3 Grp_i
\end{align*}
\]

\[
\begin{align*}
\nu_{0i} & \sim \mathcal{NID}(0, \sigma_v^2) \\
\varepsilon_{ij} & \sim \mathcal{LID}(0, \pi^2/3)
\end{align*}
\]
Random Intercept Logistic Model *in terms of probability*

- Not linear in terms of probability

\[
Pr(Y_{ij} = 1) = \frac{1}{1 + \exp \left[ - (\beta_0 + \beta_1 G_i + \beta_2 T_j + \beta_3 (G_i \times T_j) + \nu_{0i}) \right]}
\]

where \( G = \text{Group} \quad T = \text{Time} \)
Random Intercept Logistic Model

in terms of log odds (logits)

- Linear in terms of log odds (logits)

\[
\log \left( \frac{Pr(Y_{ij} = 1)}{1 - Pr(Y_{ij} = 1)} \right) = \beta_0 + \beta_1 G_i + \beta_2 T_j + \beta_3 (G_i \times T_j) + \nu_{0i}
\]
Random Intercept and Trend Model

**Within-subjects model - level 1** \((j = 1, \ldots, n_i)\)

Latent response

\[ y_{ij} = b_{0i} + b_{1i} Time_{ij} + \varepsilon_{ij} \]

**Between-subjects model - level 2** \((i = 1, \ldots, N)\)

\[ b_{0i} = \beta_0 + \beta_2 Grp_i + \nu_{0i} \]

\[ b_{1i} = \beta_1 + \beta_3 Grp_i + \nu_{1i} \]

\[
\begin{bmatrix}
\nu_{0i} \\
\nu_{1i}
\end{bmatrix}
\sim \mathcal{NID}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{\nu_0}^2 & \sigma_{\nu_0\nu_1} \\ \sigma_{\nu_0\nu_1} & \sigma_{\nu_1}^2 \end{bmatrix}\right)
\]

\[ \varepsilon_{ij} \sim \mathcal{LID}(0, \pi^2/3) \]
Mixed-effects regression model for latent response strength $y_{ij}$

$$y_{ij} = x'_{ij} \beta + z'_{ij} \nu_i + \varepsilon_{ij}$$

$i = 1 \ldots N$ subjects; $j = 1 \ldots n_i$ observations within subject $i$

$y_{ij}$ = latent response strength of observation $j$ within subject $i$

$x_{ij} = (p + 1) \times 1$ covariate vector

$\beta = (p + 1) \times 1$ vector of fixed regression parameters

$z_{ij} = r \times 1$ design vector for the random effects

$\nu_i = r \times 1$ vector of random effects for subject $i \sim \mathcal{NID}(0, \Sigma_v)$

$\varepsilon_{ij} = \text{residuals} \sim \mathcal{NID}(0, 1)$ for probit,

or $\sim \mathcal{LID}(0, \pi^2/3)$ for logistic
With model assumptions

$$\mathcal{E}(y_i) = X_i \beta$$

$$\mathcal{V}(y_i) = Z_i \Sigma_v Z_i' + \sigma^2 I_i$$

- For a random-intercepts model
  $$\mathcal{V}(y_i) = \sigma_v^2 1_i 1_i' + \sigma^2 I_i$$
  \(\Rightarrow\) compound-symmetry structure

- For more general random-effects models, \(\Rightarrow\) more general structure for \(\mathcal{V}(y_i)\)

- For probit formulation, \(y_i \sim\) multivariate normal
Notice, without $\mathbf{v}_i$

$$y_{ij} = x'_{ij}\mathbf{\beta} + \varepsilon_{ij}$$

$$\mathcal{E}(y_i) = X_i\mathbf{\beta}$$

$$\mathcal{V}(y_i) = \sigma^2 I_i$$

$\Rightarrow \mathbf{\beta}$ from MRM are not on the same scale as from a model without $\mathbf{v}_i$
Treatment-Related Change Across Time

NIMH Schizophrenia collaborative study on treatment related changes in overall severity (IMPS item # 79). Item 79, *Severity of Illness*, was scored as:

1 = normal, 2 = borderline mentally ill, 3 = mildly ill, 4 = moderately ill, 5 = markedly ill, 6 = severely ill, 7 = among the most extremely ill

The experimental design and corresponding sample sizes:

<table>
<thead>
<tr>
<th>Group</th>
<th>Sample size at Week</th>
<th>Completers</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLC (n=108)</td>
<td>107 105 5 87 2 2 70</td>
<td>65%</td>
</tr>
<tr>
<td>DRUG (n=329)</td>
<td>327 321 9 287 9 7 265</td>
<td>81%</td>
</tr>
</tbody>
</table>

*Drug = Chlorpromazine, Fluphenazine, or Thioridazine*

Main question of interest:

- Was there differential improvement for the drug groups relative to the control group?
### Descriptive Statistics

**Observed proportions ≥ “moderately ill”**

<table>
<thead>
<tr>
<th></th>
<th>week 0</th>
<th>week 1</th>
<th>week 3</th>
<th>week 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>placebo</td>
<td>.98</td>
<td>.91</td>
<td>.89</td>
<td>.71</td>
</tr>
<tr>
<td>drug</td>
<td>.99</td>
<td>.82</td>
<td>.66</td>
<td>.42</td>
</tr>
</tbody>
</table>

**Observed odds ≥ “moderately ill”**

<table>
<thead>
<tr>
<th></th>
<th>week 0</th>
<th>week 1</th>
<th>week 3</th>
<th>week 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>placebo</td>
<td>52.5</td>
<td>9.50</td>
<td>7.70</td>
<td>2.50</td>
</tr>
<tr>
<td>drug</td>
<td>80.8</td>
<td>4.63</td>
<td>1.93</td>
<td>.73</td>
</tr>
<tr>
<td>ratio</td>
<td>.65</td>
<td>2.05</td>
<td>3.99</td>
<td>3.42</td>
</tr>
</tbody>
</table>

**Observed log odds ≥ “moderately ill”**

<table>
<thead>
<tr>
<th></th>
<th>week 0</th>
<th>week 1</th>
<th>week 3</th>
<th>week 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>placebo</td>
<td>3.96</td>
<td>2.25</td>
<td>2.04</td>
<td>.92</td>
</tr>
<tr>
<td>drug</td>
<td>4.39</td>
<td>1.53</td>
<td>.66</td>
<td>-.31</td>
</tr>
<tr>
<td>difference</td>
<td>-.43</td>
<td>.72</td>
<td>1.38</td>
<td>1.23</td>
</tr>
<tr>
<td>exp (odds ratio)</td>
<td>.65</td>
<td>2.05</td>
<td>3.99</td>
<td>3.42</td>
</tr>
</tbody>
</table>
Observed Proportions across Time by Condition

- model is not linear in terms of probabilities
Observed Logits across Time by Condition
NIMH Schizophrenia Study - Severity of Illness (N = 437)  
Logistic Regression ML Estimates - *Fixed effects model*

<table>
<thead>
<tr>
<th></th>
<th>estimates</th>
<th>se</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>3.702</td>
<td>0.441</td>
<td>8.39</td>
<td>.001</td>
</tr>
<tr>
<td>Drug (0 = plc; 1 = drug)</td>
<td>-0.405</td>
<td>0.483</td>
<td>-0.84</td>
<td>.41</td>
</tr>
<tr>
<td>Time (sqrt week)</td>
<td>-1.112</td>
<td>0.233</td>
<td>-4.78</td>
<td>.001</td>
</tr>
<tr>
<td>Drug by Time</td>
<td>-0.418</td>
<td>0.256</td>
<td>-1.64</td>
<td>.11</td>
</tr>
</tbody>
</table>

\[-2 \log L = 1362.06\]

*ok if data were cross-sectional longitudinal or if \( \sigma_v = 0 \)*
Fitted Logits across Time by Condition

fixed-effects logistic regression model

\[ \log \left[ \frac{Pr(Y_{ij} = 1)}{1 - Pr(Y_{ij} = 1)} \right] = 3.70 - .41 D_i - 1.11 T_j - .42 (D_i \times T_j) \]
Fitted Proportions across Time by Condition

*fixed-effects logistic regression model*

\[
Pr(Y_{ij} = 1) = \frac{1}{1 + \exp\left[-\left(3.70 - .41 D_i - 1.11 T_j - .42 D_i T_j\right)\right]}
\]
Within-Subjects / Between-Subjects components

**Within-subjects model - level 1** \((j = 1, \ldots, n_i \text{ obs})\)

\[
\text{logit}_{ij} = b_{0i} + b_{1i}\sqrt{\text{Week}_j}
\]

**Between-subjects model - level 2** \((i = 1, \ldots, N \text{ subjects})\)

\[
b_{0i} = \beta_0 + \beta_2 \text{Grp}_i + \nu_{0i}
\]

\[
b_{1i} = \beta_1 + \beta_3 \text{Grp}_i
\]

\[
\nu_{0i} \sim \mathcal{NID}(0, \sigma_v^2)
\]
NIMH Schizophrenia Study - Severity of Illness (N = 437)

Logistic ML Estimates (se) - *random-intercepts model*

<table>
<thead>
<tr>
<th>estimates</th>
<th>se</th>
<th>z</th>
<th>p &lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>5.387</td>
<td>0.631</td>
<td>8.54</td>
</tr>
<tr>
<td>Drug (0 = plc; 1 = drug)</td>
<td>-0.025</td>
<td>0.654</td>
<td>-0.04</td>
</tr>
<tr>
<td>Time (sqrt week)</td>
<td>-1.500</td>
<td>0.291</td>
<td>-5.16</td>
</tr>
<tr>
<td>Drug by Time</td>
<td>-1.015</td>
<td>0.334</td>
<td>-3.04</td>
</tr>
<tr>
<td>Intercept variance</td>
<td>4.478</td>
<td>0.947</td>
<td></td>
</tr>
</tbody>
</table>

*Intra-person correlation* = \( 4.478 / (4.478 + \pi^2 / 3) \) = .58

\[-2 \log L = 1249.73 \quad \chi^2_1 = 112.33\]
Estimated (subject-specific) Logits across Time by Condition: \textit{random-intercepts model}

\[
\log \left[ \frac{Pr(Y_{ij} = 1)}{1 - Pr(Y_{ij} = 1)} \right] = 5.39 - 0.03 D_i - 1.50 T_j - 1.01 (D_i \times T_j) + \nu_{0i}
\]

\[
\nu_{0i} \sim \mathcal{NID}(0, \hat{\sigma}^2_{\nu} = 4.48)
\]

\(\hat{\beta}\) assesses change in (conditional) logit due to \(x\) for subjects with the same value of \(\nu_{0i}\)
Random-intercepts Logistic Regression

\[ \text{logit}_{ij} = \mathbf{x}^\prime_{ij} \boldsymbol{\beta} + \nu_{0i} \]

- every subject has their own propensity for response ($\nu_{0i}$)
- the influence of covariates $\mathbf{x}$ is determined controlling (or adjusting) for the subject effect
- the covariance structure, or dependency, of the repeated observations is explicitly modeled
\( \beta_0 \) = log odds of response for a typical subject with \( x = 0 \) and \( \nu_{0i} = 0 \)

\( \beta \) = log odds ratio for response associated with unit changes in \( x \) for the same subject value \( \nu_{0i} \)

* referred to as “subject-specific”

* how a subject’s response probability depends on \( x \)

\( \sigma^2_v \) = degree of heterogeneity across subjects in the probability of response not attributable to \( x \)

- most useful when the objective is to make inference about subjects rather than the population average

- interest is in the heterogeneity of subjects
Estimated (subject-specific) probabilities across time

Random intercepts model - placebo group

\[
P(Y_{ij} = 1) = \frac{1}{1 + \exp[-(5.39 - 0.03 D_i - 1.50 T_j - 1.01 D_i T_j + \hat{v}_{0i})]}\]
Estimated (subject-specific) probabilities across time

Random intercepts model - drug group

\[ P(Y_{ij} = 1) = \frac{1}{1 + \exp[-(5.39 - .03 D_i - 1.50 T_j - 1.01 D_i T_j + \hat{\upsilon}_{0i})]} \]
Estimated Subject-Specific Probabilities

random-intercepts logistic regression model

\[
Pr(Y_{ij} = 1) = \frac{1}{1 + \exp \left[ - (5.39 - .03 D_i - 1.50 T_j - 1.01 D_i T_j + \nu_{0i}) \right]}
\]

where \( \nu_{0i} = \begin{cases} 
-1\sigma_v & \text{and } \hat{\sigma}_v = 2.12 \\
1\sigma_v & 
\end{cases} \)
Model fit of observed marginal proportions

1. \( \hat{y}_i = X_i \hat{\beta} \)

2. calculate marginalization factor

\[
\hat{s} = \sqrt{\hat{d}} = \sqrt{\left(\hat{\sigma}_v^2 + \sigma^2\right)/\sigma^2} = \sqrt{\hat{\sigma}_v^2/\sigma^2 + 1}
\]

- \( \sigma = 1 \) for probit or \( \sigma = \pi/\sqrt{3} \) for logistic
- \( \hat{d} \) is the design effect in the sampling literature

3. marginalize \( \hat{z}_i = \hat{y}_i / \hat{s} \)

4. \( \hat{p}_i = \Phi(\hat{z}_i) \) for probit and \( \hat{p}_i = \Psi(\hat{z}_i) \) for logistic, \( \Phi \) represents the normal cdf and \( \Psi \) the logistic cdf, i.e., \( 1/[1 + \exp(-\hat{z}_i)] \)
notes:

• In practice, for logistic, \((15\pi)/(16\sqrt{3})\) works better than \(\pi/\sqrt{3}\) as \(\sigma\) (Zeger et al., 1988, Biometrics)

• Logistic is approximate; relies on cumulative Gaussian approximation to the logistic function

• For multiple random effects, calculate marginalization vector

\[
\hat{s} = \frac{1}{\sigma} \left[ \text{Diag}(\hat{V}(y_i)) \right]^{1/2}
\]

\[
- \hat{V}(y_i) = Z_i \hat{\Sigma}_\nu Z'_i + \sigma^2 I_i
\]

- \(Z_i\) = design matrix for random effects

and perform element-wise division

\[
\hat{z}_i = \hat{y}_i / \hat{s}
\]
Estimated Marginal Logits and Probabilities
SAS NLMIXED code: SCHZBINL.SAS (at website as Example 9.1)

DATA one; INFILE 'c:\mixdemo\schizx1.dat';  
INPUT id imps79 imps79b imps79o int tx week sweek txswk ;

/* get rid of observations with missing values */
IF imps79 > -9;

PROC FORMAT;
VALUE imps79b 0 = 'le mild' 1 = 'ge moderate';
VALUE tx 0 = 'placebo' 1 = 'drug';

/* fixed-effects logistic regression model */
PROC LOGISTIC DESCENDING;
MODEL imps79b = tx sweek tx*sweek;
RUN;

/* random intercept logistic regression via GLIMMIX */
PROC GLIMMIX DATA=one METHOD=QUAD(QPOINTS=21) NOCLPRINT;
CLASS id;
MODEL imps79b(DESC) = tx sweek tx*sweek / SOLUTION DIST=BINARY LINK=LOGIT;
RANDOM INTERCEPT / SUBJECT=id;
RUN;
/* random intercept logistic regression via NLMIXED */
PROC NLMIXED DATA=one QPOINTS=21;
PARMS b0=3.70 b1=-.40 b2=-1.11 b3=-.42 varu=1;
z = b0 + b1*tx + b2*sweek + b3*tx*sweek + u;
IF (imps79b=1) THEN
  p = 1 / (1 + EXP(-z));
ELSE
  p = 1 - (1 / (1 + EXP(-z)));
ll = LOG(p);
MODEL imps79b ~ GENERAL(ll);
RANDOM u ~ NORMAL(0, varu) SUBJECT=id;
ESTIMATE 'icc' varu/(((ATAN(1)*4)**2)/3)+varu);
RUN;
SAS IML code: SCHZBFIT1.SAS (at website as Example 9.2)

TITLE1 'nimh schizophrenia data - estimated marginal probabilities';
PROC IML;
/* Results from nlmixed analysis: random intercept model */;

/* covariate matrices for placebo and drug groups */;
x0 = { 1 0 0.00000 0,
       1 0 1.00000 0,
       1 0 1.73205 0,
       1 0 2.44949 0};
x1 = { 1 1 0.00000 0.00000,
       1 1 1.00000 1.00000,
       1 1 1.73205 1.73205,
       1 1 2.44949 2.44949};

/* nlmixed estimates of covariate effects and random effect variance */;
beta = {5.387, -0.025, -1.500, -1.015};
varu = {4.478};

/* marginalization of person-specific estimates */;
pi = ATAN(1)*4;
nt = 4;
ivec = J(nt,1,1);
zvec = J(nt,1,1);
evec = (15/16)**2 * (pi**2)/3 * ivec;
/* nt by nt matrix with evec on the diagonal and zeros elsewhere */;
emat = DIAG(evec);

/* variance-covariance matrix of underlying latent variable */;
vary = zvec * varu * T(zvec) + emat;

/* marginalization factor */;
sdy = SQRT(VECDIAG(vary) / VECDIAG(emat));

z0 = (x0*beta) / sdy ;
z1 = (x1*beta) / sdy;

grp0 = 1 / ( 1 + EXP(0 - z0));
grp1 = 1 / ( 1 + EXP(0 - z1));

print 'random intercept model';
print 'marginalization of person-specific estimates';
print 'marginal prob for group 0 - response' grp0 [FORMAT=8.4];
print 'marginal prob for group 1 - response' grp1 [FORMAT=8.4];
Random intercept and trend model
within-subjects / between-subjects components

within-subjects model - level 1 \((j = 1, \ldots, n_i \text{ obs})\)

\[
\text{logit}_{ij} = b_{0i} + b_{1i}\sqrt{\text{Week}_j}
\]

between-subjects model - level 2 \((i = 1, \ldots, N \text{ subjects})\)

\[
b_{0i} = \beta_0 + \beta_2\text{Grp}_i + \upsilon_{0i}
\]

\[
b_{1i} = \beta_1 + \beta_3\text{Grp}_i + \upsilon_{1i}
\]

\[
\upsilon_i \sim \mathcal{NID}(0, \Sigma_v)
\]
### Logistic ML Estimates (se) - random intercept and trend model

<table>
<thead>
<tr>
<th></th>
<th>estimates</th>
<th>se</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>5.928</td>
<td>0.948</td>
<td>6.25</td>
<td>.001</td>
</tr>
<tr>
<td>Drug (0 = plc; 1 = drug)</td>
<td>0.287</td>
<td>0.742</td>
<td>0.39</td>
<td>.70</td>
</tr>
<tr>
<td>Time (sqrt week)</td>
<td>-1.399</td>
<td>0.476</td>
<td>-2.94</td>
<td>.004</td>
</tr>
<tr>
<td>Drug by Time</td>
<td>-1.615</td>
<td>0.481</td>
<td>-3.36</td>
<td>.001</td>
</tr>
</tbody>
</table>

### Variance-covariance terms

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept var</td>
<td>6.975</td>
<td>2.908</td>
<td></td>
</tr>
<tr>
<td>Int-Time covar</td>
<td>-2.111</td>
<td>1.210</td>
<td></td>
</tr>
<tr>
<td>Time var</td>
<td>3.096</td>
<td>1.161</td>
<td></td>
</tr>
</tbody>
</table>

\[ -2 \log L = 1227.38, \quad \chi^2_2 = 21.95, \quad p < .001 \]
Estimated (subject-specific) probabilities across time

Random intercepts and trends model - placebo group

\[ P(Y_{ij} = 1) = \frac{1}{1 + \exp\left[-(5.93 + .29 D_i - 1.40 T_j - 1.62 D_iT_j + \hat{v}_0 + \hat{v}_1 T_j)\right]} \]
Estimated (subject-specific) probabilities across time

Random intercepts and trends model - drug group

\[
P(Y_{ij} = 1) = \frac{1}{1 + \exp[-(5.93 + .29 D_i - 1.40 T_j - 1.62 D_i T_j + \hat{v}_{0i} + \hat{v}_{1i} T_j)]}
\]
Estimated Marginal Logits and Probabilities
**SAS NLMIXED code: random-trend logistic regression**

*included in SCHZBINL.SAS syntax file (at website as Example 9.1)*

/* random trend logistic regression via GLIMMIX */
PROC GLIMMIX DATA=one METHOD=QUAD(QPOINTS=11) NOCLPRINT;
CLASS id;
MODEL imps79b DESC = tx sweek tx*sweek / SOLUTION DIST=BINARY LINK=LOGIT;
RANDOM INTERCEPT sweek / SUBJECT=id TYPE=UN GCORR SOLUTION;
ODS LISTING EXCLUDE SOLUTIONR; ODS OUTPUT SOLUTIONR=ebest2;
RUN;

/* logistic random-trend model via NLMixed */
PROC NLMIXED DATA=one QPOINTS=11;
PARMS b0=5.39 b1=-0.03 b2=-1.50 b3=-1.02 v0=4.48 c01=0 v1=1;
z = b0 + b1*tx + b2*sweek + b3*tx*sweek + u0 + u1*sweek;
IF (imps79b=1) THEN
  p = 1 / (1 + EXP(-z));
ELSE
  p = 1 - (1 / (1 + EXP(-z)));
ll = LOG(p);
MODEL imps79b ~ GENERAL(ll);
RANDOM u0 u1 ~ NORMAL([0,0], [v0,c01,v1]) SUBJECT=id OUT=ebest2b;
ESTIMATE ’re corr’ c01/SQRT(v0*v1);
RUN;
SAS IML code: SCHZBFIT2.SAS (at website as Example 9.3)

```
TITLE1 'nimh schizophrenia Data - estimated marginal probabilities';
PROC IML;
/* results from nlmixed analysis: random intercept & trend model */;

/* covariate matrices for placebo and drug groups */;
x0 = { 1 0 0.00000 0,
     1 0 1.00000 0,
     1 0 1.73205 0,
     1 0 2.44949 0};
x1 = { 1 1 0.00000 0.00000,
     1 1 1.00000 1.00000,
     1 1 1.73205 1.73205,
     1 1 2.44949 2.44949};

/* nlmixed estimates of covariate effects and random effect variance-covariance matrix */;
beta = { 5.928, 0.287, -1.399, -1.615};
varu = {6.975 -2.111,
       -2.111 3.096};

/* marginalization of person-specific estimates */;
pi = ATAN(1)*4;
nt = 4;
ivec = J(nt,1,1);
zmat = {1 0.00000,
        1 1.00000,
        1 1.73205,
        1 2.44949};
```

evec = (15/16)**2 * (pi**2)/3 * ivec;

/* nt by nt matrix with evec on the diagonal and zeros elsewhere */;
emat = DIAG(evec);

/* variance-covariance matrix of underlying latent variable */;
vary = zmat * varu * T(zmat) + emat;

/* marginalization factor */;
sdy = SQRT(VECDIAG(vary) / VECDIAG(emat));

z0 = (x0*beta) / sdy;
z1 = (x1*beta) / sdy;

grp0 = 1 / ( 1 + EXP(0 - z0));
grp1 = 1 / ( 1 + EXP(0 - z1));

print 'random intercept and trend model';
print 'marginalization of person-specific estimates';
print 'marginal response probability for group 0' grp0 [FORMAT=8.4];
print 'marginal response probability for group 1' grp1 [FORMAT=8.4];
Logistic GEE as marginal model

\[
\text{logit}_{ij} = x'_{ij} \beta
\]

- Working correlation of repeated observations
  - exchangeable (all are equal), AR(1), banded (\(m\)-dependent), unstructured
- robust standard errors
- does not include any subject-specific (random) effects, does not focus on heterogeneity
  \(\beta_0 = \) log odds of response among sub-population with \(x = 0\)
  \(\beta = \) log odds ratio for response associated with unit changes in \(x\) in the population of subjects
- \(\exp(\beta) = \) ratio of population frequencies
  - referred to as “population-averaged”
NIMH Schizophrenia Study - Severity of Illness (N = 437)
Logistic Regression GEE - exchangeable correlation structure

<table>
<thead>
<tr>
<th></th>
<th>GEE estimates</th>
<th>se</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>3.661</td>
<td>0.485</td>
<td>7.54</td>
<td>.001</td>
</tr>
<tr>
<td>Drug (0 = plc; 1 = drug)</td>
<td>-0.381</td>
<td>0.521</td>
<td>-0.73</td>
<td>.46</td>
</tr>
<tr>
<td>Time (sqrt week)</td>
<td>-1.094</td>
<td>0.252</td>
<td>-4.35</td>
<td>.001</td>
</tr>
<tr>
<td>Drug by Time</td>
<td>-0.449</td>
<td>0.269</td>
<td>-1.67</td>
<td>.10</td>
</tr>
</tbody>
</table>

- non-significant drug by time interaction
- working corr based on data from 7 timepts (weeks 0 to 6)
- several have little data (wks 2, 4, 5) & wk 0 is near-constant
- very poorly estimated working correlation matrix
- analysis of 4 primary timepts and UN working corr yields significant interaction (p < .047)
Estimated Marginal Logits and Probabilities
SAS GENMOD code: GEE logistic regression - SCHZGEE.SAS (at website Example 9.4)

```sas
DATA one; INFILE 'c:\mixdemo\schizx1.dat';
INPUT id imps79 imps79b imps79o int tx week sweek txswk;

/* get rid of observations with missing values */
IF imps79 > -9;

/* get rid of weeks with very few observations */
IF week EQ 0 or week EQ 1 OR week EQ 3 OR week EQ 6;

PROC FORMAT;
VALUE imps79b 0 = 'le mild' 1 = 'ge moderate';
VALUE tx 0 = 'placebo' 1 = 'drug';

/* gee logistic regression model: unstructured */
PROC GENMOD DESCENDING;
CLASS id week;
MODEL imps79b = tx sweek txswk / LINK=LOGIT DIST=BIN;
REPEATED SUBJECT=id / WITHIN=week CORRW TYPE=UN;
RUN;
```
Conclusions - mixed-effects logistic regression models useful for incomplete longitudinal dichotomous data

• can handle subjects measured incompletely or at different timepoints (missing data assumed MAR)

• degree of within-subjects variation on dichotomous outcome is important to consider (might have 3-timepoint study where 90% of subjects have same response across timepoints)

• subject-specific (or conditional) interpretation of regression coefficients

• generalizations to other categorical outcomes
  – ordinal outcomes - mixed-effects ordinal logistic regression
    * proportional odds model
    * partial or non-proportional odds model
  – nominal outcomes - mixed-effects nominal logistic regression