# Comparing the Estimates of Mixed Models for Binary or Ordinal Data

October 20, 2006

Meeting of the Society of Multivariate Experimental Psychology

Daniel J. Bauer

L.L. Thurstone Psychometric Laboratory

The University of North Carolina at Chapel Hill

## Mixed Models

- **Mixed models provide a model-based way to account for dependence in clustered or longitudinal data.**
- **Dependence is usually accommodated through the introduction of random effects.**
  - **Random intercepts to account for mean differences between units.**

## The Linear Mixed Model

- **The general form of the linear mixed model is**

$$y_{ij} = \mathbf{x}_{ij}'\boldsymbol{\beta} + \mathbf{z}_{ij}'\mathbf{u}_j + r_{ij}$$

  - **observation $i$ is nested within unit $j$.**
  - $\mathbf{x}_{ij}$ **is a vector of predictors with fixed effects.**
  - $\boldsymbol{\beta}$ **is the vector of fixed effects.**
  - $\mathbf{z}_{ij}$ **is a vector of predictors with random effects.**
  - $\mathbf{u}_j$ **is the vector of random effects for unit $j$.**
  - $r_{ij}$ **is the observation-specific residual.**
- **Typically assume that**

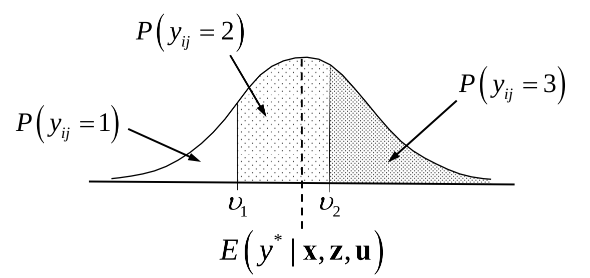
$$\begin{pmatrix}\mathbf{u}\\ \mathbf{r}\end{pmatrix} \sim N\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{u}} & 0\\ 0 & \sigma_r^2\mathbf{I}_n\end{bmatrix}\right)$$

## Mixed Model for Binary and Ordinal Outcomes

- **Assume that $y$ represents a coarse version of a continuous underlying variable $y^*$**

$$y_{ij}^* = \mathbf{x}_{ij}'\boldsymbol{\beta} + \mathbf{z}_{ij}'\mathbf{u}_j + r_{ij}$$

$$y_{ij} = c \text{ if } \upsilon^{c-1} < y_{ij}^* \le \upsilon^c \qquad (\upsilon_0 \equiv -\infty;\ \upsilon_C \equiv \infty)$$



## Scaling for Binary and Ordinal Outcomes

- **Because $y^*$ is not directly observed, its location and scale are arbitrary.**
- **Set the location by fixing first threshold: $\upsilon_1 \equiv 0$**
- **Set the scale by fixing $\sigma_r^2$**
  - **In the probit model, assume $r_{ij} \sim N(0,1)$**
  - **In the logit model, assume $r_{ij} \sim \text{logistic}\left(0, \frac{\pi^2}{3}\right)$**

## Comparing Sequential Models

- **Can we compare the estimates of sequential models?**
  - **Model 1: $y_{ij}^* = \beta_0 + u_j + r_{ij}$**
  - **Model 2: $y_{ij}^* = \beta_0 + \beta_1 x_{1ij} + u_j + r_{ij}$**
  - **Model 3: $y_{ij}^* = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_j + r_{ij}$**
- **Across these models, $\sigma_r^2$ remains constant.**
  - **In the probit model: $\sigma_r^2 = 1$**
  - **In the logit model: $\sigma_r^2 = \pi^2 / 3$**
- **Variance of $y^*$ must then change between models, putting estimates on different scales.**

## Placing Estimates on a Commensurate Scale

- **To compare the estimates of simple logistic or probit models, Winship & Mare (1983, 1984) recommended equating the model-implied marginal variance of $y^*$ across the fitted models.**

## Placing Estimates on a Commensurate Scale

- **The model-implied marginal variance of $y^*$ in the mixed model is**

$$V\left(y_{ij}^*\right) = \boldsymbol{\beta}'\boldsymbol{\Sigma}_{\mathbf{x}}\boldsymbol{\beta} + \boldsymbol{\mu}_{\mathbf{z}}'\boldsymbol{\Sigma}_{\mathbf{u}}\boldsymbol{\mu}_{\mathbf{z}} + VEC\left(\boldsymbol{\Sigma}_{\mathbf{z}}\right)' VEC\left(\boldsymbol{\Sigma}_{\mathbf{u}}\right) + \sigma_r^2$$

- **It then follows that**

$$\left(\boldsymbol{\beta}'\boldsymbol{\Sigma}_{\mathbf{x}}\boldsymbol{\beta} + \boldsymbol{\mu}_{\mathbf{z}}'\boldsymbol{\Sigma}_{\mathbf{u}}\boldsymbol{\mu}_{\mathbf{z}} + VEC\left(\boldsymbol{\Sigma}_{\mathbf{z}}\right)' VEC\left(\boldsymbol{\Sigma}_{\mathbf{u}}\right) + \sigma_r^2\right)^{-1} V\left(y_{ij}^*\right) = 1$$

- **To rescale the implied variance to equal $a$, note that**

$$a\left(\boldsymbol{\beta}'\boldsymbol{\Sigma}_{\mathbf{x}}\boldsymbol{\beta} + \boldsymbol{\mu}_{\mathbf{z}}'\boldsymbol{\Sigma}_{\mathbf{u}}\boldsymbol{\mu}_{\mathbf{z}} + VEC\left(\boldsymbol{\Sigma}_{\mathbf{z}}\right)' VEC\left(\boldsymbol{\Sigma}_{\mathbf{u}}\right) + \sigma_r^2\right)^{-1} V\left(y_{ij}^*\right) = a$$

## Placing Estimates on a Commensurate Scale

- **Recall that** $s^2 V(y_{ij}^*) = V(sy_{ij}^*)$**, where** $s$ **is a scaling factor**

$$\underbrace{a\left(\boldsymbol{\beta}'\boldsymbol{\Sigma}_{\mathbf{x}}\boldsymbol{\beta} + \boldsymbol{\mu}_{\mathbf{z}}'\boldsymbol{\Sigma}_{\mathbf{u}}\boldsymbol{\mu}_{\mathbf{z}} + VEC\left(\boldsymbol{\Sigma}_{\mathbf{z}}\right)' VEC\left(\boldsymbol{\Sigma}_{\mathbf{u}}\right) + \sigma_r^2\right)^{-1}}_{s^2} V\left(y_{ij}^*\right) = a$$

- **We can apply this to the mixed model equation as follows:**

$$s = a^{1/2}\left(\boldsymbol{\beta}'\boldsymbol{\Sigma}_{\mathbf{x}}\boldsymbol{\beta} + \boldsymbol{\mu}_{\mathbf{z}}'\boldsymbol{\Sigma}_{\mathbf{u}}\boldsymbol{\mu}_{\mathbf{z}} + VEC\left(\boldsymbol{\Sigma}_{\mathbf{z}}\right)' VEC\left(\boldsymbol{\Sigma}_{\mathbf{u}}\right) + \sigma_r^2\right)^{-1/2}$$

$$sy_{ij}^* = \mathbf{x}_{ij}'\left(s\boldsymbol{\beta}\right) + \mathbf{z}_{ij}'\left(s\mathbf{u}_j\right) + \left(sr_{ij}\right)$$

$$VAR\left(sy_{ij}^*\right) = a$$

## Comparing Fitted Models

- **Given the equation**

$$sy_{ij}^* = \mathbf{x}_{ij}'\left(s\boldsymbol{\beta}\right) + \mathbf{z}_{ij}'\left(s\mathbf{u}_j\right) + \left(sr_{ij}\right)$$

**the rescaled fixed effects are** $s\boldsymbol{\beta}$

**the rescaled covariance parameters are** $s^2\boldsymbol{\Sigma}_{\mathbf{u}}$ **and** $s^2\sigma_r^2$

- **For any two models to be compared, can rescale estimates using a common** $a$ **value so that the units are commensurate.**

## Simulated Data Example

- **10,000 Clusters, 1 to 10 observations per cluster**
- **Population Model**

$$y_{ij}^* = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_j + r_{ij} \qquad y = 1 \text{ if } y_{ij}^* > 0;\ \text{else } y = 0$$

$$\beta_0 = 0;\ \beta_1 = \beta_2 = 1$$

$$\sigma_u^2 = .05;\ \sigma_r^2 = .15;\ r_{ij} \sim N$$

$$\boldsymbol{\mu}_{\mathbf{x}} = \mathbf{0}$$

$$\boldsymbol{\Sigma}_{\mathbf{x}} = \begin{pmatrix} .1+.3 & 0 \\ 0 & .1+.3 \end{pmatrix}$$

$CORR(x_1, x_2) = 0$

## Sequence of Models

- **Fit two mixed-effects probit models to** $y$
  - **Model 1:** $P(y_{ij} = 1) = g\left(\beta_0 + \beta_1 x_{1ij} + u_j\right)$
  - **Model 2:** $P(y_{ij} = 1) = g\left(\beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_j\right)$

  $g^{-1}$ **chosen to be probit link function.**

- **For comparison, fit two linear mixed models to simulated** $y^*$
  - **Model 1:** $y_{ij}^* = \beta_0 + \beta_1 x_{1ij} + u_j + r_{ij}$
  - **Model 2:** $y_{ij}^* = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_j + r_{ij}$

*Estimates for models fit to simulated data.*

| *Effects* | *Fit to $y^*$* Linear | *Fit to y* Probit | Rescaled (a=1) | Rescaled (a=2) |
|---|---|---|---|---|
| *Model 1* | | | | |
| $\beta_0$ | 0.01 | 0.01 | 0.01 | 0.01 |
| $\beta_1$ | 1.02 | 1.52 | 1.03 | 1.45 |
| $\sigma_u^2$ | 0.15 | 0.29 | 0.13 | 0.26 |
| $\sigma_r^2$ | 0.44 | 1.00 | 0.46 | 0.91 |
| *Model 2* | | | | |
| $\beta_0$ | 0.00 | 0.01 | 0.00 | 0.01 |
| $\beta_1$ | 1.00 | 2.61 | 1.00 | 1.42 |
| $\beta_2$ | 0.99 | 2.62 | 1.00 | 1.42 |
| $\sigma_u^2$ | 0.05 | 0.30 | 0.04 | 0.09 |
| $\sigma_r^2$ | 0.15 | 1.00 | 0.15 | 0.29 |

## Conclusions

- **For continuous outcomes, often recommended to fit and compare sequentially fit models.**
- **For binary and ordinal outcomes, sequentially fit models cannot be meaningfully compared unless the estimates are rescaled to equate the model-implied variance of $y^*$.**

## Of Additional Interest…

- **Attention to scaling also clarifies several other issues that arise with mixed models for binary/ordinal outcomes:**
  - **Why marginal and conditional model estimates differ (inclusion of random effects changes scale)**
  - **Why estimators (e.g., MQL, PQL) that produce biased variance component estimates also produce biased fixed effects estimates (on wrong scale)**
  - **Why misspecification of the variance component structure can bias the fixed effects estimates (on wrong scale)**