Practice in Growth Curve Modeling for Women's Menstrual Cycle - Fundamental Overview -

Chel Hee Lee¹  Angela Baerwald²

¹Clinical Research Supporting Unit
College of Medicine
University of Saskatchewan

²Department of Obstetrics, Gynecology and Reproductive Sciences
College of Medicine
University of Saskatchewan

Saskatoon SAS Users Group Meeting
September 16, 2015
1. Introduction
   - Facing Problems
   - Linear Mixed-effect Model

2. Parameteric Approach to Modeling Curves
   - Useful Models
   - Useful Techniques

3. Trial and Error
   - 1st trial: Oh No! ...
   - 2nd trial: Making a progress ...
   - 3rd trial: It's not done yet ...
   - 4th trial: Much better now ...

4. Questions
“Essentially, all models are wrong, but some are useful.”

(Box and Draper, 1986, p. 424)

- It is a best guess.
- It is an approximation.
- It would be a translation of observation into a model.
Hormones Involved in Women’s Menstrual Cycle

Subject-specific Trajectory (in Logarithm Scale)
Linear Mixed-Effect Models (Singer, 1998)

Level 1 (within person)

\[ Y_{ij} = \pi_{0j} + \pi_{1j}(\text{TIME})_{ij} + r_{ij}, \quad \text{where } r_{ij} \sim N(0, \sigma^2) \]

Level 2 (between-person)

\[ \pi_{0j} = \beta_{00} + u_{0j}, \quad \text{where } \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix} \right) \]

\[ \pi_{1j} = \beta_{10} + u_{1j}, \]

Unconditional Linear Growth Curve

\[ Y_{ij} = \underbrace{[\beta_{00} + \beta_{10} \text{TIME}_{ij}]}_{\text{FIXED}} + \underbrace{[u_{0j} + u_{1j} \text{TIME}_{ij} + r_{ij}]}_{\text{RANDOM}} \]

Note that a 3-level model can be made if individuals within groups are tracked over time.
Using PROC MIXED in SAS

Commands for a linear time trend model

```sas
PROC MIXED DATA = datafile;
   CLASS id time;
   MODEL y = time / SOLUTION CHISQ;
   REPEATED time / TYPE=UN SUBJECT=id;
RUN;
```

Commands for a subject-specific model

```sas
PROC MIXED DATA = datafile;
   CLASS id;
   MODEL y = time / SOLUTION CHISQ;
   RANDOM intercept time / TYPE=UN SUBJECT=id;
RUN;
```
Concerns Bothering Us

We are working with ...

- A small size of participants with many repeated measurements,
Concerns Bothering Us

We are working with ...

- A small size of participants with many repeated measurements,
- An unbalanced data set observed at unequally spaced times, and
Concerns Bothering Us

We are working with ...  
- A small size of participants with many repeated measurements,  
- An unbalanced data set observed at unequally spaced times, and  
- An incomplete data due to missing values.
Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
- Addressing a difference between groups,
Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
- Addressing a difference between groups,
- Dealing with missing data and dropout,
Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
- Addressing a difference between groups,
- Dealing with missing data and dropout,
- Binning data or using data as it is,
We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
- Addressing a difference between groups,
- Dealing with missing data and dropout,
- Binning data or using data as it is,
- Centralizing, normalizing, orthogonalizing data, and
Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
- Addressing a difference between groups,
- Dealing with missing data and dropout,
- Binning data or using data as it is,
- Centralizing, normalizing, orthogonalizing data, and
- Applying either or both models,
Concerns Bothering Us

We are faced with difficulties of ...

- Determining the level of a model of interest,
- Incorporating a person-level covariate,
- Addressing a difference between groups,
- Dealing with missing data and dropout,
- Binning data or using data as it is,
- Centralizing, normalizing, orthogonalizing data, and
- Applying either or both models,
- Choosing an appropriate pattern of covariance structure, etc.
Parameteric Approaches to Modeling Curves

Nonlinear Model

\[ y_i = u(x_i) + \epsilon_i, \quad i = 1, 2, \ldots, n \]

- \( p \)-th Degree Polynomial Regression Model (Johnson et al., 2013)

\[ u(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_p x_i^p \]
Parameteric Approaches to Modeling Curves

Nonlinear Model

\[ y_i = u(x_i) + \epsilon_i, \quad i = 1, 2, \ldots, n \]

- *p*-th Degree Polynomial Regression Model (Johnson et al., 2013)
  \[ u(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_p x_i^p \]

- Exponential Growth Curve Model (Zwietering M H et al., 1990)
  \[ u(x_i) = \beta_0 e^{\beta_1 x_i} \]
Parameteric Approaches to Modeling Curves

Nonlinear Model

\[ y_i = u(x_i) + \epsilon_i, \quad i = 1, 2, \ldots, n \]

- \( p \)-th Degree Polynomial Regression Model (Johnson et al., 2013)
  \[ u(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_p x_i^p \]

- Exponential Growth Curve Model (Zwietering M H et al., 1990)
  \[ u(x_i) = \beta_0 e^{\beta_1 x_i} \]

- Trigonometric Model (Cornelissen Germaine, 2014)
  \[ u(x_i) = \mu + A \cos(wx_i + \phi) \]
Modeling Techniques

Piecewise Modeling Approach (Naumova et al., 2001)

\[ u(x_i) = \beta_0 + \beta_1 x_i + \beta_2 (x_i - c)\delta, \]

where \( \delta = 1 \) if \( x_i > c \) and \( \delta = 0 \) if \( x_i \leq c \).

Cubic-Splines Modeling Approach (Gurrin et al., 2005)

\[ u(x; \theta) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{i=1}^{m} \theta_i (x - c_i)^3_+ \]
**OH No! This is not what I wanted. (Problems?)**

**3rd degree polynomial regression**

![Graph showing fitted values over time](image.png)
OH No! This is not what I wanted. (Problems?)

3rd degree polynomial regression
Oh! It’s better! However, ... Disadvantages???

Piecewise-4th degree polynomial regression
Oh! It’s better! However, ... Disadvantages???

Piecewise-4th degree polynomial regression
Trial and Error

4th trial: Much better now ...

Oh! It’s better! However, ... Disadvantages???

Piecewise-4th degree polynomial regression

Chel Hee Lee, Angela Baerwald (U of S)
Let’s do more work! Still, ... Disadvantages???

3rd degree polynomial regression with trigonometric functions
Let’s do more work! Still, ... Disadvantages???

3rd degree polynomial regression with trigonometric functions
Much Better Now! Disadvantages???

Natural Cubic Splines
Much Better Now! Disadvantages???

Natural Cubic Splines

**Trial and Error**

4th trial: Much better now ...

Chel Hee Lee, Angela Baerwald (U of S)

Practice in Growth Curve Modeling

2015-09-16 14 / 17
QUESTIONS?
References I


