<u>Topics for these notes</u>:

- *Linear mixed models with a random intercept, one sample data*
- *Repeated measures ANOVA*
- *Polynomial contrasts for time*

<u>Associated reading</u>: Sections 1, 2.1 in 'LMM introduction…' course notes.

1 Introduction

Here is some associated reading for linear mixed models:

- *Linear Mixed Models for Longitudinal Data*; Verbeke and Molenberghs; Springer; 2000.

- *Longitudinal Data Analysis*, Hedeker and Gibbons, Wiley, 2006, Chapters 4-7.

- *Applied Longitudinal Analysis,* Fitzmaurice, Laird and Ware, Wiley, 2011, Chapters 8-9.

- SAS Help and Documentation (version 9.1, 9.2, 9.3).

- Both SAS and R are used to fit linear mixed models in these notes.

- Linear mixed model methods are the state-of-the-art methods for analyzing clustered data (including longitudinal data) when the outcome variable is continuous and approximately normal.

- If the outcome is not ‘approximately normal’, one of the following might be considered (these will be discussed later):

    - Transform the outcome

    - Mixture distribution model

    - Non-normal outcome model

- A *mixed model* may have random as well as fixed effects.

- Mixed models allow for more complicated error covariance structures, unlike simpler general linear models.

- Parameters usually estimated using Maximum likelihood (ML) or restricted maximum likelihood (REML) methods.

- Random effects are estimated using empirical Bayes methods.

- Tests for fixed-effect parameters in the model are usually conducted using functions of estimated parameters that have exact or approximate t or F statistics.

- I make the effort to distinguish linear mixed model (LMM) methods from the LMMs themselves, since there are other methods that can be used to fit LMMs from what is considered standard LMM methodology.

    - A simple mixed model with a random intercept and other fixed effects could be fit using what is referred to as repeated measures ANOVA.
    - A model without random effects but with a non-independent error covariance structure (that could be considered as a special case of a mixed model) could be fit using generalized least squares.
    - ‘Multivariate GLM’ methods such as MANOVA could be used to fit an LMM with no random effects but unstructured error covariance structure.

- Differences in results obtained between using these alternative approaches and standard LMM methods are often minor, and in some cases will be the same. However, these alternative approaches have their limitations in that they can only be applied to specific types of mixed models.

- Loose definition of a *linear mixed model*: a regression-type model with fixed effects, plus either random effects or a non-simple error covariance structure, or both.

- Loose definition of a *linear mixed model method*: a contemporary inferential method that can be applied to an LMM.

Example 1: biomarker or complement levels in the body as outcome (see first set of notes); time as predictor; random terms for subjects (intercept and time slope); time can be modeled as a continuous variable since time point measurements were unequally spaced.

Example 2: growth curve data, where the outcome variable is height; the main fixed-effect predictor is age; random effects for subjects to allow differences over time (splines could be used for both fixed-effect and random-effects for age); covariates may also be included, such as subject's gender, race.

Example 3: Families are selected to participate in a survey regarding health insurance. Each member of the family will be included in the study. The outcome variable may be insurance cost. Random effects for family and subject within family can be included.

Example 4: arm length and leg length growth are measured for subjects once a year for 10 years, and then modeled with a linear mixed model. The outcome is length; there is a predictor variable indicating whether the measure is on the arm or leg; there is a time variable; a random intercept term for subjects would be the simplest term to distinguish subjects, more complex terms could be added and tested. The 'Kronecker' covariance structure can be employed to deal with these 'doubly repeated measures'.

2 Linear mixed models with a random intercept

- To help introduce linear mixed models, we will first consider a special case which I will call a *random intercept model*.
  - This is basically a general linear model with an additional random effect called a random intercept.
  - This random intercept can be defined for any cluster unit, but here we consider it for subjects.
  - This model offers one simple way to account for longitudinal data.

- These notes contain the following: data analysis for random intercept models, fundamentals of writing contrasts for time and *group***time* effects, and discussion of how current LMM methods are related to an older longitudinal data analysis technique called repeated measures ANOVA.

- Considering longitudinal studies, a random intercept term for subjects will account for between-subject variability, and will also induce a correlation structure for the responses.
  - This structure is very simple and often not the best for longitudinal data, but generally it is an improvement over no correlation structure at all.
  - For one-sample repeated measures (such as the Ramus data), this model will allow us to estimate the *intraclass correlation* (ICC) which is the proportion of total variability in the data (not including those that can be accounted for by the fixed effects) that is due to between-subject differences.

- The basic model is

$$Y_{ij} = \beta_0 + \beta_1 x_{1ij} + \ldots + \beta_{p-1} x_{p-1,ij} + b_i + \varepsilon_{ij}$$
$$= \mathbf{X}_{ij}\boldsymbol{\beta} + b_i + \varepsilon_{ij} \; ,$$

where Y is the outcome, $\mathbf{X}_{ij} = (x_1,\ldots,x_{p-1})$ is a row vector of predictors, both for subject i at time j, and where $\varepsilon_{ij} \sim N(0,\sigma_\varepsilon^2)$ and $b_i \sim N(0,\sigma_b^2)$. These random terms are assumed to be independent of each other. The main element that distinguishes this from a general linear model is the addition of the random term b_i. We also use subject and time indices here, with the addition of the repeated measures.

2.1 One-sample data

2.1.1 Time as class

2.1.1.1 Fit using LMM methods

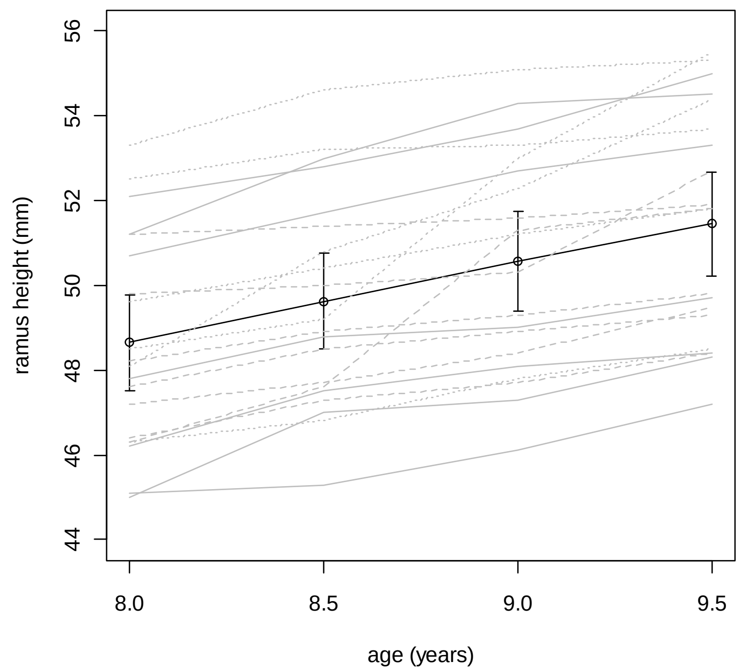
- Recall the Ramus data, where the Ramus bone in the jaw was measured on many boys at four fixed ages: 8, 8½, 9 and 9½ years of age. The random intercept model for these data is

$$Y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{3ij} + b_i + \varepsilon_{ij}$$

where i=1,…, n subjects, and j=1,…, r times; x_{pij}=1 for p=j, 0 otherwise, for p=1,…,3 are the dummy variables for time; the highest level of time is used as the reference level.

- This is the full-rank model since there are no linear dependencies in $\mathbf{X}$. Note that the i indices could be dropped from the x terms (x_{ij} replaced with x_j) since they are subject invariant.

Ramus data



The comparable less-than-full-rank model is

$$Y_{ij} = \mu + \tau_j + b_i + \varepsilon_{ij}, \tag{1}$$

where τ_j contains the effects for the 4 times.

- Fitting the full-rank and less-than-full-rank models is equivalent if we use SAS's approach to solving the generalized inverse (which will 'drop' the column associated with the highest level of time when computing the generalized inverse).

- The model implies the following. Can you verify?
    - $Cov(Y_{ij}, Y_{ij'}) = \sigma_b^2$ for $j \neq j'$
    - $Cov(Y_{i'j}, Y_{ij}) = 0$ for $i \neq i'$
    - $Var(Y_{ij}) = \sigma_b^2 + \sigma_\varepsilon^2$.

Note 1: The covariance structure for repeated measures within an individual is called *compound symmetric*. The correlation is assumed to be the same between any two time points, whether they are close together in time or further apart in time.

Note 2: For types of clustered data other than longitudinal data, the CS structure may be intuitive. For example, a sample may be taken from an individual (e.g., blood or urine sample), split into subsamples, each of which is given different treatments, after which some measurement is made.

- Before taking time to explain the mechanics and associated inferential techniques in detail, let's just look at the basic model fit using REML estimation. Note that all parameters (the two variance parameters and two fixed effect parameters) are fit simultaneously.

```
proc mixed data=long.ramus_uni;
  class boy age;
  model height = age / solution;
  random intercept / subject=boy;
  lsmeans age; run;
```

Dimensions

Covariance Parameters 2

Columns in X 5

Columns in Z Per Subject 1

Subjects 20

Max Obs Per Subject 4

Dependent Variable	*height*
Subject Effect	*boy*
Estimation Method	*REML*
Residual Variance Method	*Profile*

Covariance Structure	*Variance Components*
Fixed Effects SE Method	*Model-Based*
Degrees of Freedom Method	*Containment*

Solution for Fixed Effects

Effect	age	Estimate		Std Error	DF	t Value	Pr>|t|
Intercept		51.45	$\hat{\mu}$	0.583	19	88.30	<.0001
age	8	-2.795	$\hat{\tau}_1$	0.265	57	-10.56	<.0001
age	8.5	-1.825	$\hat{\tau}_2$	0.265	57	-6.89	<.0001
age	9	-0.880	$\hat{\tau}_3$	0.265	57	-3.32	0.0016
age	9.5	0	$\hat{\tau}_4$	.	.	.	.

```
Number of Observations Used 80

Covariance Parameter Estimates

Cov Parm Subject Estimate

Intercept boy 6.0896
Residual    0.7009
Fit Statistics
-2 Res Log Likelihood 268.6
AIC (smaller is better) 272.6
```

$\hat{\sigma}_b^2$ (Intercept boy estimate), $\hat{\sigma}_\varepsilon^2$ (Residual estimate)

```
Type 3 Tests of Fixed Effects

       Num  Den
Effect  DF    F Value   Pr > F
age      3   57   41.42  <.0001


Least Squares Means
                           Std
Effect age  Estimate   Error  DF t Value Pr>|t|
age      8    48.655  0.583  57  83.50  <.0001
age    8.5    49.625  0.583  57  85.17  <.0001
age      9    50.570  0.583  57  86.79  <.0001
age    9.5    51.450  0.583  57  88.30  <.0001
```

- Note that the ML estimates of Ramus bone size at each age, $\hat{\mu}+\hat{\tau}_j$, j=1,…,4, turn out to be the sample means at each age. These are the 'least squares means.'

Fit using R: Here is the same model fit using R software. Note that you need to install the nlme package first; the library statement will then load the nlme package.

```
library(nlme)
ramus= read.table("C:/strand_folders/…ramus_uni.csv", header = T, sep
= ",",skip = 0)
results<-lme(height~factor(age),random=~1|boy,data=ramus)
results

> results
Linear mixed-effects model fit by REML
  Data: ramus
  Log-restricted-likelihood: -134.3059
  Fixed: height ~ factor(age)
   (Intercept)     factor(age)8.5     factor(age)9     factor(age)9.5
        48.655      0.970               1.915            2.795

Random effects:
 Formula: ~1 | boy     (Intercept)    Residual
  StdDev:              2.467705       0.837226
```

- There are some key differences in displayed default output for SAS PROC MIXED and the LME function in R:

  - The ‘factor’ function is used to indicate that age will be modeled as a class variable; the lowest level of age is set as the reference level (rather than the highest that is the default in SAS). Thus, the estimates above are for age 8 (intercept), and then for ages 8½, 9 and 9½ relative to age 8.

  - The log (restricted) likelihood is shown, rather than the -2 log likelihood.

  - Standard deviation of estimated variance components are shown, rather than variances.

- Despite these differences, you will find that the model fit is actually the same. R performs adequately for simpler linear mixed models. In a later set of notes, a comparison of R and SAS is given for slightly more complex models. For much more complicated models, R may not perform as well as SAS, although if you wait long enough, this may change (as R updates seem to come weekly).

2.1.1.2 Repeated measures ANOVA

- Repeated measures (RM) ANOVA is an older method of analyzing longitudinal data.

  - It extends inferential methods for the univariate GLM in order to account for correlated measurements.

  - These extensions allow one to conduct tests for fixed-effect and random-effect terms in model (1), and estimate the (within-subject error and between-subject variance components).

- Here are the basic steps for RM ANOVA:
  - (a) compute the ANOVA table for each source of variation in the model, including the random intercept term, using standard methods;
  - (b) calculate associated expected mean square quantities for each source;
  - (c) use the expected mean square quantities to construct F-tests for terms in the model;
  - (d) estimate variance components using MOM and the MS quantities from the ANOVA table.

- For model (1), we will generally get the same or similar results whether using the RM ANOVA approach, or using standard LMM methods. Most of these details are provided in the Classical Methods notes, however some will be highlighted in this section.

- Unlike the LMM methods, RM ANOVA can only be applied to very specific mixed models such as (1). However, one benefit of viewing the calculations this way is that we can see how sums-of-squares are partitioned and construct related tests. LMM methods do not use ANOVA tables, but build statistics based on estimated parameters via ML estimation or a related estimation method called REML.

<u>ANOVA table for one-sample data (balanced case)</u>. Note: a subject-time interaction term could also be included in the model. However, in that case the interaction SS is confounded with the residual SS; tests for subject effects can only be carried out assuming no interaction.

Source	DF	SS	MS	E(MS)
Subjects	n–1	$r\sum(\bar{Y}_{i\bullet}-\bar{Y}_{\bullet\bullet})^2$	$SS_S/(n-1)$	$\sigma_\varepsilon^2+r\sigma_b^2$
Time	r–1	$n\sum(\bar{Y}_{\bullet j}-\bar{Y}_{\bullet\bullet})^2$	$SS_T/(r-1)$	$\sigma_\varepsilon^2+\frac{n}{(r-1)}\sum(\tau_j-\bar{\tau}_\bullet)^2$ "Q(Time)"
Resid.	$(n-1)(r-1)$	$\sum\sum(Y_{ij}-\bar{Y}_{i\bullet}-\bar{Y}_{\bullet j}+\bar{Y}_{\bullet\bullet})^2$	$SS_R/[(n-1)(r-1)]$	σ_ε^2
Total	nr–1	$\sum\sum(\bar{Y}_{ij}-\bar{Y}_{\bullet\bullet})^2$		

- Expected mean square quantities can be derived using standard theory approaches and employing assumptions from the model.

  - F-tests can be constructed for Time and Subject sources by E(MS) quantities.

  - Notice that $E(MS_T) = \sigma_\varepsilon^2 + Q(T)$ and $E(MS_R) = \sigma_\varepsilon^2$. The difference between these two quantities is $Q(T)$, which is 0 when $\tau_1 = \tau_2 = \ldots = \tau_r = 0$ and is greater than zero otherwise.

  - Thus, an F-test can be constructed using MST/ MSR in order to test for time effects.

  - Similarly, a test for subject effects can be constructed as MSS / MSR, however since this is a variance component, it is testing H_0: $\sigma_b^2 = 0$.

- Using the method of moments, we can estimate variance components and the intra-class correlation (ICC).

Variance components

$$\mathrm{MS_S} = \hat{\sigma}_\varepsilon^2 + r\hat{\sigma}_b^2,$$

$$\mathrm{MS_R} = \hat{\sigma}_\varepsilon^2$$

$$\Rightarrow \hat{\sigma}_\varepsilon^2 = \mathrm{MS_R},$$

$$\hat{\sigma}_b^2 = (\mathrm{MS_S} - \mathrm{MS_R})/r$$

ICC

$$\hat{\sigma}_b^2 / (\hat{\sigma}_b^2 + \hat{\sigma}_\varepsilon^2) = [\mathrm{MS_S} - \mathrm{MS_R}] / [\mathrm{MS_S} + (r-1)\,\mathrm{MS_R}]$$

- Here is the PROC GLM code for the Ramus data analysis.

```
proc glm order=data data=ramus;  class boy age;
 model height=boy age / solution;  random boy /test;
 contrast 't2 versus t1'  age -1 1 0 0;
 contrast 't3 versus t1' age -1 0 1 0;
 contrast 't4 versus t1'  age -1 0 0 1;
 contrast 'linear' age -3 -1  1 3;
 contrast 'quadratic'  age  1 -1 -1 1;
 contrast 'cubic' age -1  3 -3 1;
 contrast 'deviations' age  1 -1 -1 1, age -1  3 -3 1; run;
```

This is an example of the model introduced in Section 2.1; “age” is the time variable (associated with fixed effects τ_j); “boy” is the random subject variable (associated with random effects b_i).

Source	DF	Type III SS	Mean Square	F Value	Pr > F
boy	19	476.1250000	25.0592105	35.75	<.0001
age	3	87.0910000	29.0303333	41.42	<.0001

SS_age is the same as SS_T defined previously.

Parameter	Estimate		Std Error	t Value	Pr > |t|
Intercept	50.625	B	0.4489	112.77	<.0001
boy 1	-0.425	B	0.5920	-0.72	0.4758
boy 2	-1.800	B	0.5920	-3.04	0.0036
. . .					
boy 20	0.000	B	.	.	.
age 8	-2.795	B	0.2648	-10.56	<.0001
age 8.5	-1.825	B	0.2648	-6.89	<.0001
age 9	-0.880	B	0.2648	-3.32	0.0016
age 9.5	0.000	B	.	.	.

The fixed intercept relates to the estimate for the last age and last boy, since boy is fitted in the model (incorrectly) as a fixed effect. Note that this is technically not part of the RM ANOVA. For PROC MIXED, the fixed intercept estimate was equivalent to the estimate at the last age (averaged over boy, 9½).

Source	Type III Expected Mean Square
boy	Var(Error) + 4 Var(boy)
age	Var(Error) + Q(age)

Note that these match the E(MS) quantities defined previously, but we have a specific form for "Q(age)" = "Q(time)".

```
Tests of Hypotheses for Mixed Model Analysis of Variance
Dependent Variable: height
                                              Mean
Source              DF    Type III SS       Square    F Value    Pr > F
boy                 19      476.125         25.059      35.75    <.0001
age                  3       87.091         29.030      41.42    <.0001
Error: MS(Error)    57       39.954          0.701
                              Contrast          Mean
Contrast            DF          SS            Square   F Value    Pr > F
t2 versus t1         1       9.4090         9.4090     13.42    0.0005
t3 versus t1         1      36.6723        36.6723     52.32    <.0001
t4 versus t1         1      78.1203        78.1203    111.45    <.0001
linear               1      87.0489        87.0489    124.19    <.0001
quadratic            1       0.0405         0.0405      0.06    0.8109
cubic                1       0.0016         0.0016      0.00    0.9621
deviations           2       0.0421         0.0211      0.03    0.9704
```

The ‘reference cell’ contrasts are not orthogonal. Thus the Contrast SS’s do not add up to SS_age circled above.

The polynomial contrasts are orthogonal. Thus the Contrast SS’s do add up to SS_age circled above.

The ‘deviations’ contrast here is equivalent to a linear lack-of-fit test. See below for

Calculating the ICC:

$$\hat{\sigma}_b^2 \,/\, (\hat{\sigma}_b^2 + \hat{\sigma}_\varepsilon^2) = [MS_S - MS_R] \,/\, [MS_S + (r-1)\, MS_R]$$

$$= [25.059 - 0.701]/[25.059+3(0.701)]$$

= 0.897 (same as with the LMM approach)

2.1.1.3 Contrasts for time – an introduction

- Usually, if the overall test for time is found to be significant, comparisons involving specific time points are of interest.
  - In terms of model (1), we are interested in contrasts of the form **C** (same as discussed near the end of the GLM review).
  - We discussed how the **C** matrix in a contrast statement does not necessarily need to have row contrasts (where coefficients sum to 1). However, here we do consider **C** matrices with such constraints.
  - The fixed parameters in model (1) can be expressed as $\boldsymbol{\beta} = (\mu\ \tau_1\ \tau_2 \ldots \tau_r)^t$. The examples below consider r=4.

- Orthogonal polynomial contrasts. There are advantages to creating orthogonal contrasts.

    - Contrasts $\sum c_i\mu_i$ and $\sum d_i\mu_i$ are orthogonal if $\sum c_i d_i = 0$ (balanced data case).

    - For a factor with r levels (r–1 degrees of freedom), we can construct r–1 orthogonal contrasts for specific tests involving the levels of the factor.

    - Orthogonality among the r–1 contrasts is a nice property because the sum of squares for the factor is partitioned into r–1 non-overlapping quantities that can be used to conduct r–1 independent tests.

    - The orthogonal contrasts can further be made orthonormal by rescaling each contrast to have unit length, if desired.

- Choosing contrast coefficients to make orthogonal polynomials
  - There are a couple of properties we want the coefficients of polynomial contrasts to have.

    (1) Coefficients of $\mathbf{c}_{it}$ add up to 0 (i.e., true contrasts).

    - Helps to ensure orthogonality (although won't solely ensure it)

    - Tests of contrasts will then be invariant to any scale changes to $\mathbf{c}_{it}$.

    (2) Measures the strength of the given polynomial trend, and when trend is absent, the measure ($\mathbf{c}_{it}\boldsymbol{\beta}$) should be 0.

- One way to get the coefficients is as follows.

  - Let **T** denote the time matrix for intercept, linear, quadratic and cubic terms (polynomial trends in separate rows) for times t=0, 1, 2,…,r. The intercept is included in the first row so that **T** is $r\times r$.

  - Let $\mathbf{S}\ \mathbf{S}^t = \mathbf{T}\ \mathbf{T}^t$, where **S** is an $r\times r$ lower triangular matrix (Cholesky factorization).

  - An orthogonalized polynomial matrix that will satisfy (1) and (2) above is $\mathbf{P} = \mathbf{S}^{-1}\ \mathbf{T}$.

  - Individual rows of **P** can be rescaled to allow for whole numbers (or rescaled by any amount, for that matter; invariance will hold).

- Illustration:

```
proc iml;
time={1 1 1 1,
      0 1 2 3,
      0 1 4 9,
      0 1 8 27};

orthopoly=inv(t(root(time*t(time))))*time;
print 'time matrix' time;
print 'orthogonalized time matrix',
  orthopoly [format=8.4]; run;
```

time matrix, T

1	1	1	1
0	1	2	3
0	1	4	9
0	1	8	27

orthogonalized time matrix, P (these are in fact what we call orthonormal, since the sum of the squares of each row is 1)

orthopoly

0.5000	0.5000	0.5000	0.5000
-0.6708	-0.2236	0.2236	0.6708
0.5000	-0.5000	-0.5000	0.5000
-0.2236	0.6708	-0.6708	0.2236

multiply row of P by ? to get whole numbers

$\times\ 2 = (1\ 1\ 1\ 1)$	intercept
$\times\sqrt{20} = (-3\ -1\ 1\ 3)$	linear
$\times\ 2 = (1\ -1\ -1\ 1)$	quadratic
$\times\sqrt{20} = (-1\ 3\ -3\ 1)$	cubic

- In this case, the r–1 orthogonal polynomials are the last 3 rows of **P**. Let C denote this matrix (without the first row):

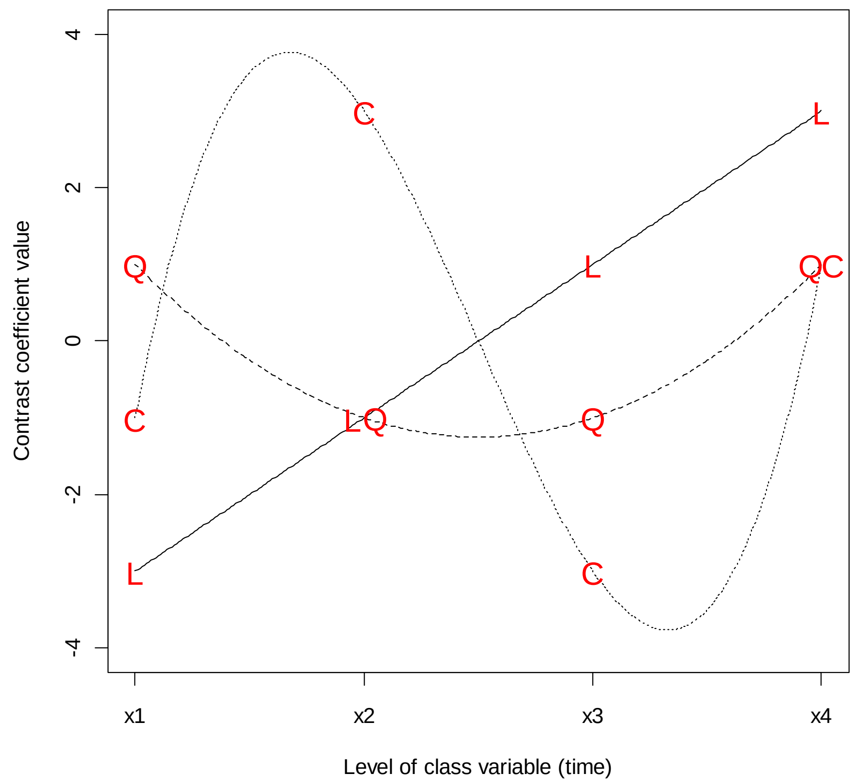
$$\mathbf{C} = \begin{pmatrix} -3 & -1 & 1 & 3 \\ 1 & -1 & -1 & 1 \\ -1 & 3 & -3 & 1 \end{pmatrix}$$

- For the effects model for the Ramus data, we would have $\boldsymbol{\beta} = (\mu\ \tau_1\ \tau_2\ \tau_3\ \tau_r)^t$ and

$$\mathbf{C} = \left(\begin{array}{c|cccc} 0 & -3 & -1 & 1 & 3 \\ 0 & 1 & -1 & -1 & 1 \\ 0 & -1 & 3 & -3 & 1 \end{array}\right)$$

- For the remainder of this section, we will discuss **C** matrices that apply to either the means model, or to the τ effects for the effects model (i.e., dropping the first column if the effects model is considered).

Visualizing polynomial coefficients



L=(-3, -1, 1, 3)

Q=(1, -1, -1, 1)

C=(-1, 3, -3, 1)

These coefficients will determine the strength of the respective polynomial trends. Here, r=4 (as with the Ramus data).

- Technically we make an implicit assumption that the levels of time are equally spaced for the contrasts to make sense. Also, the best choice of polynomial coefficients depends on r. E.g., you could choose (-1, 0, 1) for a linear contrast and (1, -2, 1) for a quadratic contrast if r=3.

- Time versus reference contrasts:

  - You pick a reference time and compare all other times to that time. For example, compare 'time 1' to each of the other times:

$$\mathbf{C} = \begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

  - The matrix is associated with the hypothesis H_0: $\tau_1 = \tau_2,\ \tau_1 = \tau_3,\ \tau_1 = \tau_4$ (i.e., all τ_i equal). Note that these contrasts are not orthogonal.

- Other types of contrasts that might be of interest include…

    - *Profile contrasts* to compare successive pairs of adjacent contrasts

    - *Helmert contrasts* to compare time point j with the average of subsequent time points, for j=1,…,r–1

    - *Deviation contrasts* that compare time point j with the average of all other time points, j=1,…,r. Note that there would be r such deviation contrasts, while the other types naturally have r–1. Deviation contrasts are associate with the question: How much does a given time point mean deviate from the average of the others?

- **Note**: Often we will define **C** as the coefficients for time effects only (i.e., drop the first column in the **C** matrices above. (Another way to think about it is to not include the intercept in the model. Either way, results will be the same.)

Contrasts with PROC MIXED

```
proc mixed data=long.ramus_uni method=ml;
 class boy age; model height= age;
 random intercept / subject=boy;
 contrast 't2 versus t1'  age -1 1 0 0;
 contrast 't3 versus t1' age -1 0 1 0;
 contrast 't4 versus t1'  age -1 0 0 1;
 contrast 'linear' age -3 -1  1 3;
 contrast 'quadratic'  age  1 -1 -1 1;
 contrast 'cubic' age -1  3 -3 1;
 contrast 'deviations' age  1 -1 -1 1,
                        age -1  3 -3 1;
run;
```

Portion of output with contrast results:

Contrasts

Label	Num DF	Den DF	F Value	Pr > F
t2 versus t1	1	57	14.13	0.0004
t3 versus t1	1	57	55.07	<.0001
t4 versus t1	1	57	117.32	<.0001
linear	1	57	130.72	<.0001
quadratic	1	57	0.06	0.8061
cubic	1	57	0.00	0.9611
deviations	2	57	0.03	0.9689

- For a factor with r–1 degrees of freedom, it is sometimes recommended that the number of contrasts constructed should have a total of r–1 degrees of freedom.

- But it is possible to do more (if justifications can be made for doing so). If this is done, however, the Contrast SS quantities will add up to more than the SS for that factor, as illustrated above.

- Since the 'linear', 'quadratic' and 'cubic' contrasts are orthogonal, their SS quantities add up to the SS for Time.

- We've also added…
  - the set of contrasts that compare times to time 1
  - the 'deviations' contrast, which answers the question: Is the simple linear model sufficient for the data, or are higher-order terms necessary? The linear term is clearly sufficient. This is a type of *lack of fit* test. Specifically, it is a linear lack of fit test, where H_0: Linear model is sufficient; H_1: Linear model is not sufficient.

- Note that for the Ramus data, there were 4 time points and thus we cannot test for trends beyond cubic.

2.1.2 Time as continuous – LMM versus RM ANOVA methods

- The previous results strongly suggest that growth of the Ramus bone is very linear, and so one could argue that a simple linear trend for time in the model is sufficient.

- We can also use Repeated Measures ANOVA in this case, for the model $Y_{ij} = \beta_0 + \beta_1 x_{ij} + b_i + \varepsilon_{ij}$. Here, x_{ij} is a continuous measure for time (e.g., x_{ij} =8½ for all i and j=2).

- As with the previous analyses that considered time as a class variable, analyses using standard LMM methods and RM ANOVA will yield the same results for these data when considering time as a continuous variable.

- For the following output, on the left is the analysis using PROC MIXED; on the right is the fit using RM ANOVA with PROC GLM, for comparison. For now, we'll only be concerned with the estimates of parameters in the mixed model and not hypothesis tests. Hence, output is abbreviated.

- Later on we will discuss the methods that are used in PROC MIXED to fit the model. Note that now I am treating *age* as a continuous variable (in the Classical notes *age* was modeled as a class variable so that various contrast statements could be included).

- For this simple model, the two approaches yield the same or similar results. Notice that unlike GLM, random effects are not included in the MODEL statement in MIXED.

```
proc mixed data=long.ramus_uni;
  class boy;
  model height = age / solution;
  random intercept / subject=boy; run;

Covariance Parameter Estimates

Cov Parm      Subject    Estimate
Intercept     boy          6.0953
Residual                   0.6779

Fit Statistics

-2 Res Log Likelihood          267.2
AIC (smaller is better)        271.2

Solution for Fixed Effects

                   Std
Effect    Estimate Error  DF   t Value   Pr>|t|
Intercept  33.78   1.55   19     21.83   <.0001
age         1.87   0.16   59     11.33   <.0001
```

```
proc glm order=data data=long.ramus_uni;
  class boy;
  model height=boy age / solution;
  random boy / test; run;

Source   DF   Type III SS       MS   F-Value Pr > F
boy      19    476.1250    25.0592     36.97 <.0001
age       1     87.0489    87.0489    128.41 <.0001

                              Standard
Parameter  Estimate           Error    t Value   Pr>|t|
Intercept    32.9225 B        1.4985     21.97   <.0001
age           1.8660          0.1647     11.33   <.0001


Tests of Hypotheses for Mixed Model Analysis of
Variance

Source   DF   Type III SS        MS F Value Pr > F
boy      19      476.1250   25.0592   36.97 <.0001
age       1      87.04890   87.0489  128.41 <.0001
Error    59      39.99610    0.6779
```

Note: when 'boy' is in the model, estimates for each boy will also be included in the Parameter Estimates, where each is followed by a 'B' (to indicate non-uniqueness); I did not include them here.

- Question: How do estimates of σ_b^2 compare between the two methods of fitting?

  Mixed: $\hat{\sigma}_b^2 = 6.095$

  GLM: $\hat{\sigma}_b^2 = (25.06 - 0.68)/4 = 6.095$ **The same!**

- With either approach, the estimated intercept and slope for age are 32.9 and 1.866, respectively.

- The expected mean square quantities in the RM ANOVA for boy and age are similar to before, although the *Q(age)* function will reflect the fact that it is continuous.

- Tests for boy and age both use the standard MSE in the denominator of the F-statistics. In this case, contrasts for time are not relevant, but estimates of growth at specific ages can be obtained. The analysis using standard LMM methods yields equivalent results and is presented later.