Mixed Effects Models

Rick White

May 12, 2015

Outline

- Linear Models (continuous response)
- Mixed Effects vs Fixed Effects models
- Examples of Mixed Effects models
- Other types of response variables
- Questions

Data setup for linear models

We assume $Y \sim N(\mu, \sigma^2)$ and the data are independent.

μ and σ^2 are parameters that describe centre and spread of the distribution respectively.

We can shift the normal distribution without changing its shape.

This allows us to split the data into 2 pieces, one fixed, the other random.

- $Y_i = \mu + \varepsilon_i$
- μ is a fixed value and ε is the random error in our data.

We assume $\varepsilon \sim N(0, \sigma^2)$

We can model both the fixed and random portions of our data.

Review of Fixed Effects Linear Models

Regression, ANOVA, ANCOVA are all fixed effects linear models.

We model μ as a function of some predictors.

Example: $\mu_{ij} = a + bx_{ij} + T_i$

x_{ij} is an observed numeric variable. a, b and T_i are fixed effects, T_i is associated with the levels of an unspecified categorical variable.

ε_{ij} is random described by a single parameter σ^2.

Approximately 95% of the errors will be between -2σ and 2σ.

We assume the errors are independent.

Estimating Fixed Effects

We do not know the true value of the fixed effects but we can estimate it using the data.

The estimates of the parameters will follow a normal distribution because of the Central Limit Theorem.

$\hat{b} \sim N(b, se^2_{\hat{b}})$

The standard error (se) associated with $\hat{b}$ is related to σ, the sample size, the spread of x_{ij} and its correlation with the other predictors.

There is a 95% chance that the true value of b is between $\hat{b} \pm 1.96 se_{\hat{b}}$.

Random effects Models

Random effects models allow us to model the variance of our data.

Example: Suppose we want to measure aptitude of students using a standard test. We need a sample of students to take the test.

- Randomly select some schools (i).
- Randomly select a set of classes within each school (j).
- Randomly select some students from each class (k).

The overall average score is μ.
The school score is $Y_i = \mu + \gamma_i$.
The class score is $Y_{ij} = \mu + \gamma_i + \delta_{ij}$.
The student score is $Y_{ijk} = \mu + \gamma_i + \delta_{ij} + \varepsilon_{ijk}$

Random Effects continued

Because we selected schools, classes and students randomly from a larger group we treat these effects as random. Since these effects are nested, they form a hierarchical linear model.

The random effects assumed to be independent with

$\gamma \sim N(0, \sigma_s^2)$ (Between school variation).
$\delta \sim N(0, \sigma_c^2)$ (Between class but within school variation).
$\varepsilon \sim N(0, \sigma^2)$ (Between student but within class variation).

This model allows us to estimate how much variation in aptitude score between students is due to the school they attend and the class they are in.

Random Effects can depend on parameters as well. Anything beyond a random intercept model is a big jump in complexity.

Random Effects continued

Random effects models assume data are conditionally independent given the random effects.

Unconditionally, observations on the same unit are correlated.

If we have $Y_{ij} = \mu + \gamma_i + \delta_{ij}$ being the jth observation on the ith subject.

$\gamma \sim N(0, \sigma_B^2)$ and $\delta \sim N(0, \sigma_W^2)$

We can compute the correlation between observations on the same and different subjects.

$Cor(Y_{1,1}, Y_{2,1}) = Cor((\gamma_1 + \delta_{11}), (\gamma_2 + \delta_{21})) = 0$

$Cor(Y_{1,1}, Y_{1,2}) = Cor((\gamma_1 + \delta_{11}), (\gamma_1 + \delta_{12})) = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2}$

Mixed Effects Models

Mixed Effects models contain both fixed and random effects.

The random effects can control for repeated measurements on the same sampled unit or cluster. This is a random intercept model.

Usually the fixed effects are the parameters of interest and can vary within any level of the random effects. They will be tested at the appropriate level within the model.

Mixed effects models do not require a balanced design which means the model can easily handle missing data.

Traditional models like repeated measures ANOVA requires the data to be balanced within the random effects. No missing data.

Example: Orthodontic measurements over time

The Data:

- distance from the pituitary to the pterygomaxillary fissure.
- Age of the subject (8, 10, 12 and 14) (within subject)
- Subject id: 27 in total
- Sex, Male (n=16) or Female (n=11) (between subject)

Lets fit distance as a function of age and subject (ignoring gender for now).

- The first model uses fixed effects only, treating subject as a blocking factor.
- The second model treats subject as a random effect.

Fixed Effects model for within subject factor

Subject is included as a fixed blocking factor.

```
## Anova Table, Response = distance
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Age        3 237.19  79.064 38.0396 2.986e-15
## Subject   26 518.38  19.938  9.5925 3.375e-15
## Residuals 78 162.12   2.078
```

```
##             Estimate Std. Error t value
## (Intercept)   21.162      0.760   27.85
## Age10          0.981      0.392    2.50
## Age12          2.463      0.392    6.28
## Age14          3.907      0.392    9.96
```

Mixed Effects model for within subject factor

Subject is included as a random effect.

```
## Anova Table, Response = distance
##     NumDF DenDF F.value    Pr(>F)
## Age     3    78   38.04 2.998e-15
```

```
##            Var    SD
## Subject  4.465 2.113
## Residual 2.078 1.442
```

```
##             Estimate Std. Error   df t value
## (Intercept)   22.185      0.492 43.4   45.07
## Age10          0.981      0.392 78.0    2.50
## Age12          2.463      0.392 78.0    6.28
## Age14          3.907      0.392 78.0    9.96
```

Comparing the results

The results from the 2 models are identical except for the error associated with the Intercept.

Age is a within subject factor so the ANOVA test for Age and the estimated effects are the same. If the design was unbalanced the results would be similar but not the same.

In the fixed effects model, we can compute the variance components using the expected means squares of the ANOVA table. This is non-trivial if the design is unbalanced.

In the mixed effects model, we get the variance components directly as a single parameter. We do not get the individual random effects directly.

Models with only between subject factors

In order to control for repeated measures, we have to average the response over subject before we can use a fixed effects model.

If we do not average, we will overstate the significance of our between subject factor.

```
## Anova Table, Response = average distance
##           Df Sum Sq Mean Sq F value   Pr(>F)
## Sex        1 35.116  35.116  9.2921 0.005375
## Residuals 25 94.479   3.779
```

```
##             Estimate Std. Error t value
## (Intercept)    24.97      0.486   51.38
## SexFemale      -2.32      0.761   -3.05
```

Mixed Effects model approach

Mixed effects model works without having to average and is accurate even if there is imbalance over the subjects.

```
## Anova Table, Response = distance
##     NumDF DenDF F.value    Pr(>F)
## Sex     1    25  9.2921 0.005375
```

```
##              Var    SD
## Subject  2.547 1.596
## Residual 4.930 2.220
```

```
##              Estimate Std. Error df t value
## (Intercept)     24.97      0.486 25   51.38
## SexFemale       -2.32      0.761 25   -3.05
```

Comparing the results

The estimated effect for sex, its standard error and significance test are the same in the 2 models.

The MSE for the average can be computed from the variance components in the mixed effects model. $3.779 = 2.547 + 4.930/4$

In the fixed effects model, we have variance of the average. In the mixed effects model we have the variance broken into its 2 components, within and between subject.

If we increase the number of observations within subject, we can reduce the variance down to a minimum of the between subject variance.

In order to seriously increase power we need to increase the number of subjects.

Within and between subject factors in the same model.

We cannot average over subject with a within subject factor in the model.

If we do not average, the fixed effects model does not properly account for Sex as a between subject factor and uses the wrong error term. This overstates the significance of Sex. The correct error term for Sex is Subject.

```
## Anova Table, Response = distance
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Age        3 237.19  79.064  38.040 2.986e-15
## Sex        1 140.46 140.465  67.581 3.513e-12
## Subject   25 377.91  15.117   7.273 6.062e-12
## Residuals 78 162.12   2.078
```

Using repeated measures ANOVA instead

This model works and is correct because we have complete data. If some subjects were not observed at all ages, this model could not be used.

```
##
## Error: Subject
##           Df Sum Sq Mean Sq F value  Pr(>F)
## Sex        1  140.5  140.46   9.292 0.00538
## Residuals 25  377.9   15.12
##
## Error: Within
##           Df Sum Sq Mean Sq F value   Pr(>F)
## Age        3  237.2   79.06   38.04 2.99e-15
## Residuals 78  162.1    2.08
```

Mixed Effects Model approach

```
## Anova Table, Response = distance
##      NumDF DenDF F.value    Pr(>F)
## Age      3    78  38.040 2.998e-15
## Sex      1    25   9.292  0.005375

##              Var    SD
## Subject  3.260 1.805
## Residual 2.078 1.442

##               Estimate Std. Error df t value
## (Intercept)    23.131      0.542 38   42.66
## Age10           0.981      0.392 78    2.50
## Age12           2.463      0.392 78    6.28
## Age14           3.907      0.392 78    9.96
## SexFemale      -2.321      0.761 25   -3.05
```

Comparing the results

The results of the repeated measures ANOVA and the mixed effects model are identical.

The Age and Sex results are the same as those obtained when we modelled them separately.

We can compute the correct test for Sex using the ANOVA table from the fixed effects model. It's just the software cannot do it automatically.

Repeated Measures ANOVA breaks the model into two pieces. If the design is unbalanced, it cannot separate the components properly.

The mixed effects model is able to compute the correct components without having to separate the pieces of the model.

Within subject factor (Unbalanced case)

Fixed Effects Model

```
## Anova Table, Response = distance
##            Df Sum Sq Mean Sq F value    Pr(>F)
## Age         3 223.94  74.646  48.577 4.676e-16
## Subject    26 414.28  15.934  10.369 6.920e-14
## Residuals 60  92.20   1.537
```

Mixed Effects Model

```
## Anova Table, Response = distance
##     NumDF  DenDF F.value    Pr(>F)
## Age     3 60.027  36.312 1.643e-13
```

Either model is fine but the choice affects the interpretation.

Between subject factors (Unbalanced case)

Fixed Effects model on average observation

```
## Anova Table, Response = average distance
##            Df  Sum Sq Mean Sq F value   Pr(>F)
## Sex         1  57.085  57.085  12.607 0.001554
## Residuals 25 113.196   4.528
```

Mixed Effects Model on raw data

```
## Anova Table, Response = distance
##     NumDF  DenDF F.value   Pr(>F)
## Sex     1 22.776  13.035 0.001489
```

Mixed effects takes variation in sample size into account. Result is more accurate.

Together in a Mixed Effects model

```
## Anova Table, Response = distance
##      NumDF  DenDF F.value    Pr(>F)
## Age      3 59.947  35.703 2.305e-13
## Sex      1 23.960  10.980  0.002918
```

```
##             Var    SD
## Subject  3.478 1.865
## Residual 1.553 1.246
```

```
##              Estimate Std. Error   df t value
## (Intercept)     23.37      0.553 35.4   42.26
## Age10            1.22      0.408 60.7    3.00
## Age12            2.35      0.369 59.8    6.36
## Age14            3.79      0.382 59.8    9.92
## SexFemale       -2.59      0.782 24.0   -3.31
```

Comparing the results for the unbalanced case

In the unbalanced case, fixed effects models give different results that the mixed effects model for both the within subject models and the between subject models.

For within subject models, a fixed effects approach means the results are only valid for this specific set of subjects.

For between subject models, fixed effects models on the average fails to adjust for the relationship between the variance of the average and the sample size .

Fitting both within and between factors together gives a different result than fitting each separately. In the unbalanced case the within factors can be partially confounded with the between factors.

A model with many fixed effects

Math Achievement Scores from 7185 students from 160 schools.

Student level predictors are

- Sex (M/F)
- Minority (Yes/No)
- SES (socio-economic status) as a numeric rating.

School level predictors are

- Size of the school in 1000's of students
- Sector (Public or Catholic)
- PRACAD (percentage students on the academic track)
- DISCLIM (numeric rating of the discrimination climate).

We fit a mixed effects model with school as a random effect (students are repeated measures within the school)

The Anova table

```
## Anova Table, Response = Math Achievement Score
##            NumDF  DenDF F.value    Pr(>F)
## Sex            1 5148.8   62.36 3.553e-15
## Minority       1 2734.6  240.77 < 2.2e-16
## SES            1 6776.5  348.18 < 2.2e-16
## Size           1  154.0   14.31  0.000221
## Sector         1  148.1    4.08  0.045275
## PRACAD         1  165.1   39.87 2.413e-09
## DISCLIM        1  160.3    4.06  0.045602
```

Note the denominator degrees of freedom (DF) are not integer valued. This is generally true mixed effects models with unbalanced data.

DF is a parameter in a statistical distribution. It is related to sample size.

The fixed and random effects

```
##              Var    SD
## School     1.462 1.209
## Residual 35.897 5.991
```

```
##                  Estimate Std. Error   df t value
## (Intercept)        10.653      0.460  180   23.17
## SexFemale          -1.253      0.159 5149   -7.90
## MinorityYes        -3.043      0.196 2735  -15.52
## SES                 1.971      0.106 6777   18.66
## Size                0.823      0.218  154    3.78
## SectorCatholic      0.802      0.397  148    2.02
## PRACAD              4.283      0.678  165    6.31
## DISCLIM            -0.375      0.186  160   -2.01
```

More than one Random Effect (Hierarchical or Crossed models)

The models above demonstrate that we can fit 2 levels of data in the same model using Mixed Effects model.

We can fit multiple levels simultaneously using this technique. Each level is defined by a random effect.

A hierarchical model requires the random effects be nested to create a hierarchy.

$Y_{ijk} = \mu + \gamma_i + \delta_{ij} + \varepsilon_{ijk}$

However random effects do not have to be nested they can be crossed.

$Y_{ijk} = \mu + \gamma_i + \delta_j + \varepsilon_{ijk}$

Example: Student evaluation of instructor's lectures

Students rate their lectures between 0 and 100. Each lecture is rated by multiple students and each student can rate multiple lectures.

The dataset has 21446 ratings from 2487 students. The lectures were taught by 318 instructors from 3 different departments. The goal is to test for a difference between departments and course type (service or not).

We need to account for repeated observations on the instructor and repeated observations by the student.

```
## ANOVA Table, Response = Score
##              NumDF  DenDF F.value    Pr(>F)
## service          1 4290.7  39.297 3.998e-10
## dept             2  350.2   4.857  0.008309
## service:dept     2 3974.8   6.905  0.001015
```

The fixed effects

The estimated group means are below:

```
##   service dept Estimate   SE  DF
## 1   FALSE    6     54.7 1.19 502
## 2    TRUE    6     50.2 1.17 453
## 3   FALSE    4     57.1 1.07 560
## 4    TRUE    4     56.4 1.14 700
## 5   FALSE    9     55.9 1.45 376
## 6    TRUE    9     50.0 1.55 465
```

The fixed effects show that the lecture ratings are similar for Dept 6 and 9 but are higher for Dept 4. In general service courses receive a lower rating within the 3 departments.

The random effects and correlations

The variance components of the model are below:

```
##                Var     SD
## Student      49.80  7.057
## Instructor   98.35  9.917
## Residual    584.59 24.178
```

Observations from difference students on different instructors are uncorrelated.

Same student on different instructors: $\rho = 0.08$

Different students on same instructor: $\rho = 0.14$

Each instructor/student combination appeared only once in the dataset so there is no same student same instructor correlation estimate.

Testing Random Effects

We can test the random effects using a likelihood ratio test.

We refit the the model without a random effects and evaluate the change in the likelihood.

```
## Likelihood ratio test for Student random effect
##    Chisq Chi Df Pr(>Chisq)
##   532.47      1  < 2.2e-16
```

```
## Likelihood ratio test for Instructor random effect
##    Chisq Chi Df Pr(>Chisq)
##   2508.9      1  < 2.2e-16
```

Mixed Effects model for non-normal responses.

For other distributions, the mean and variance are related and cannot be separated

The link function relates the fixed effects to the mean and transforms the mean to the entire real line. We also put the random effects within the link function.

Dist	Link	Variance	Dist	Link	Variance
Poisson	$\log(\mu)$	μ	NB	$\log(\mu)$	$\mu + \mu^2/r$
Binomial	$\text{logit}(\mu)$	$\mu(1-\mu)$	NB	$\log(\mu)$	$\tau^2\mu$

$g(\mu_{ij}) = \alpha + \beta x_{ij} + \gamma_i$ where $\gamma_i \sim N(0, \sigma^2)$.

Example: Contagious bovine pleuropneumonia (CBPP)

We have repeated measures on herd of zebu cattle in Ethiopia. There are 15 herds with 3 or 4 repeated measures. The response is the number of cattle with CBPP.

```
## Likelihood Ratio Test
##          Df    LRT    Pr(Chi)
## period   3 25.61 1.151e-05
```

```
##          Var     SD
## herd 0.4123 0.6421
```

```
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.398      0.231   -6.05 1.47e-09
## period2       -0.992      0.303   -3.27 1.07e-03
## period3       -1.128      0.323   -3.49 4.74e-04
## period4       -1.580      0.422   -3.74 1.82e-04
```

Questions?

www.stat.ubc.ca/SCARL

STAT 551 - Stat grad students taking this course offer free statistical advice. Fall semester every academic year.

SOS Program - An hour of free consulting to UBC graduate students. Funded by the Provost and VP Research.

Short Term Consulting Service - Advice from Stat grad students. Fee-for-service on small projects (less than 15 hours).

Hourly Projects - SCARL professional staff. Fee-for-service consulting.

The End