Mixed Effects Models

Rick White

May 18, 2016

Outline

- Linear Models
- Mixed Effects models
- Other types of 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.

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

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 $\hat{b} \pm 1.96se_{\hat{b}}$.

Random effects Models

Random effects models allow us to model the variance of our data in a hierarchical way.

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. The assumptions for the random effects are

$\gamma \sim N(0, \sigma_s^2)$ (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).

All random effects are independent.

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.

Mixed Effects Models

Mixed Effects models combine both random and fixed effects into one model.

Usually the fixed effects are the parameters of greatest interest in the model.

The random effects serve to control for repeated measurements on the same sampled unit or within the same cluster.

The fixed effects can exist at any level of the hierarchy and are 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. This is a huge advantage over traditional repeated measures ANOVA models.

Add fixed effect to our school example

Assume $\mu_{ijk} = \mu + A_i + B_{ij} + C_{ijk}$, where

- ▶ A is the province where the school is located
- ▶ B is the time of the class (Morning or Afternoon)
- ▶ C is the gender of the student

Province varies between schools, Time varies within school but between class, and Gender varies within class but between student.

Our mixed effects model is
$$Y_{ijk} = \mu + A_i + B_{ij} + C_{ijk} + \gamma_i + \delta_{ij} + \varepsilon_{ijk}$$

Gender is tested at the student level, Time at the class level and Province at the school level.

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

```
## 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)   22.185      0.277   79.96
## 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

```
## 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.

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
```

Both within and between subject factors in the model.

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

If we do not average, the fixed effects model doesn't realize that Sex is a between subject factor and uses the wrong error term. This will overstate 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
```

Within subject factor if data is not balanced

Fixed Effects Model

```
## Anova Table, Response = distance
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Age        3 259.28  86.428 34.7130 3.888e-13
## Subject   26 443.74  17.067  6.8548 4.227e-10
## Residuals 60 149.39   2.490
```

Mixed Effects Model

```
## Anova Table, Response = distance
##     NumDF  DenDF F.value     Pr(>F)
## Age     3 62.078  28.733 8.985e-12
```

Either method is fine but it affects the interpretation of results.

Imbalance for between subject factors

Fixed Effects model on average observation

```
## Anova Table, Response = distance
##           Df  Sum Sq Mean Sq F value   Pr(>F)
## Sex        1  45.136  45.136  10.639 0.003192
## Residuals 25 106.059   4.242
```

Mixed Effects Model on raw data

```
## Anova Table, Response = distance
##     NumDF  DenDF F.value   Pr(>F)
## Sex     1 24.679  10.347 0.003605
```

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 62.442 28.3719 1.071e-11
## Sex     1 25.248  9.1667  0.005612


##             Var    SD
## Subject   3.132 1.770
## Residual  2.477 1.574


##             Estimate Std. Error   df t value
## (Intercept)   23.203      0.582 44.4   39.86
## Age10          0.738      0.489 62.4    1.51
## Age12          2.575      0.478 63.0    5.39
## Age14          3.946      0.473 61.6    8.34
## SexFemale     -2.352      0.777 25.2   -3.03
```

More on Mixed Effects models

Mixed effects models assume data are independent only if we
condition on the random effects.

Unconditionally, observations on the same subject 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.

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

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

More than 1 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 Instructors

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 distributions other than the normal model, we cannot separate the mean from the variance. The two are related.

We also have the link function to contend with.

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	$\log(\mu)$	$\tau^2\mu$

In this case we put the random effects within the link function.

$g(\mu_{ij}) = \alpha + \beta x_{ij} + \gamma_i$ 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