Random Effects and Mixed Models

David M. Rocke

May 25, 2021

ANOVA—Fixed and Random Effects

- We will review the analysis of variance (ANOVA) and then move to random and fixed effects models
- Nested models are used to look at levels of variability (days within subjects, replicate measurements within days)
- Crossed models are often used when there are both fixed and random effects.
- These can be used with binary and count responses as well as numerical responses.

Software

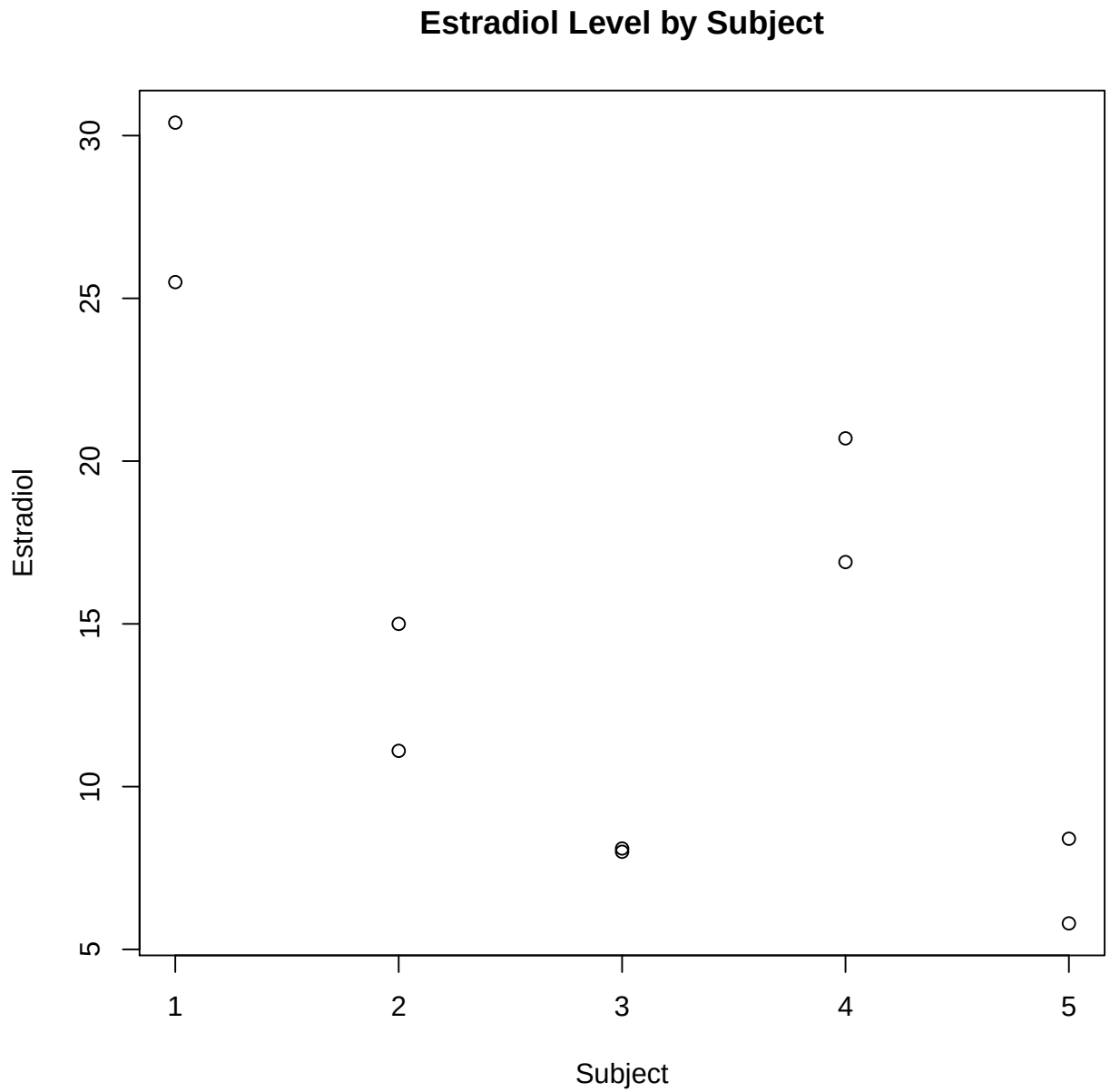
- SAS uses PROC GLM or PROC MIXED for numerical responses, and PROC GLIMMIX for numerical/binary responses.
- R uses lmer and glmer in the package lme4.
- These models can be complex and difficult to understand
- But are widely used in epidemiology, especially for longitudinal data and data clustered by hospital, herd, litter, etc.

Fixed and Random Effects

- A fixed effect is a factor that can be duplicated at a later time (dosage of a drug)
- A random effect is one that cannot be duplicated
  - Patient/subject
  - Repeated measurement
- There can be important differences in the analysis of data with random effects
- The error term is always a random effect

Endocrine data from Rosner

- 5 subjects from the Nurses' Health Study
- One blood sample each
- Each sample assayed twice for estradiol (and three other hormones)
- The within-subject variability is strictly technical/assay
- Variability within a person over time will be much greater


Estradiol Level by Subject
Estradiol
30
25
20
15
10
5
1
2
3
4
5
Subject

Fixed Effects One-Way Anova

For subject $i = 1, 2 \ldots, k$ and replicate $j = 1, 2, \ldots m$,

$$
\begin{aligned}
y_{ij} &= \mu_i + \epsilon_{ij} \\
y_{ij} &= \mu + \alpha_i + \epsilon_{ij} \\
\alpha_i &= \mu_i - \mu \\
\sum_{i=1}^{k} \alpha_i &= 0 \qquad \text{(This is not the parametrization used by R)} \\
\epsilon_{ij} &\sim N(0, \sigma^2_\epsilon)
\end{aligned}
$$

Fixed Effects One-Way Anova

$$
\begin{aligned}
E(MSE) &= \sigma_\epsilon^2 \\
E(MSA) &= Q(\alpha_1, \ldots, \alpha_k) + \sigma_\epsilon^2 \\
H_0: &\quad Q(\alpha_1, \ldots, \alpha_k) = 0 \\
H_0: &\quad \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0 \quad \text{equivalently} \\
MSA/MSE &\sim F(k-1, k(m-1)) \quad \text{under the null}
\end{aligned}
$$

($n = km$). All these statistics are printed out by default by `lm` which assumes a fixed effects model.

Random Effects One-Way Anova

For subject $i = 1, 2 \ldots, k$ and replicate $j = 1, 2, \ldots m$,

$$
\begin{aligned}
y_{ij} &= \mu + \alpha_i + \epsilon_{ij} \\
\epsilon_{ij} &\sim N(0, \sigma^2_\epsilon) \\
\alpha_i &\sim N(0, \sigma^2_\alpha) \\
E(MSE) &= \sigma^2_\epsilon \\
E(MSA) &= m\sigma^2_\alpha + \sigma^2_\epsilon \\
H_0 : & \quad \sigma^2_\alpha = 0 \\
MSA/MSE &\sim F(k-1, k(m-1)) \qquad \text{under the null} \\
\hat{\sigma}^2_\alpha &= (MSA - MSE)/m
\end{aligned}
$$

- This is called a method-of-moments estimator because it depends only on expected values of the mean squares.
- We usually use more sophisticated methods, but this one makes sense.
- If the number of replicates is not the same, this is harder to use and requires some method of determining the expected mean squares.
- SAS PROC GLM can do this, but we usually use the fancier methods.
- Note that, in this case, the hypothesis test is the same for fixed and random effects models.

Estradiol Data Analysis

```
> anova(lm(Estradiol ~ Subject,data=endocrin))
Analysis of Variance Table

Response: Estradiol
          Df Sum Sq Mean Sq F value   Pr(>F)
Subject    4 593.31 148.329  24.546 0.001747 **
Residuals  5  30.21   6.043
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Replication error variance is 6.043, so the standard deviation of replicates is 2.46 pg/mL
Compare this to average levels across subjects from 8.05 to 18.80
Estimated variance across subjects is $(148.329 - 6.043)/2 = 71.143$
Standard deviation across subjects is 8.43 pg/mL

If we average the replicates, we get five values, the standard deviation of which is also 8.43

Estradiol Data Analysis

Replication error variance is 6.043, so the standard deviation of replicates is 2.46 pg/mL
Estimated variance across subjects is $(148.329 - 6.043)/2 = 71.143$
Standard deviation across subjects is 8.43 pg/mL

Model below is intercept + random intercept per subject

```
> summary(lmer(Estradiol ~ 1+(1|Subject),data=endocrin))

Scaled residuals:
    Min      1Q  Median      3Q     Max
-0.8254 -0.6972 -0.1150  0.6703  1.2114

Random effects:
 Groups   Name        Variance Std.Dev.
 Subject  (Intercept) 71.143   8.435
 Residual              6.043   2.458
Number of obs: 10, groups:  Subject, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)   14.990      3.851   3.892
```

Fasting Blood Glucose

- Part of a larger study that also examined glucose tolerance during pregnancy
- Here we have 53 subjects with 6 tests each at intervals of at least a year
- The response is glucose as mg/100mL

Fasting Blood Glucose Analysis

```
> anova(lm(FG ~ Subject,data=fg2))
Analysis of Variance Table

Response: FG
           Df Sum Sq Mean Sq F value    Pr(>F)
Subject    52  10936 210.310  2.9235 9.717e-09 ***
Residuals 265  19064  71.938
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 >

Estimated within-Subject variance is 71.938,
   so the standard deviation is 8.48 mg/100mL
Estimated between-Subject variance is (210.310 - 71.938)/6  =  23.062
Estimated between-Subject sd = 4.80 mg/100mL
The variance of the 53 means is 35.05, which is larger than 23.062
   because it includes a component of the within-subject variance.
```

Nested Random Effects Models

- Cooperative trial with 6 laboratories, one analyte (7 in the full data set), 3 batches per lab (a month apart), and 2 replicates per batch
- Estimate the variance components due to labs, batches, and replicates
- Test for significance if possible
- Effects are lab, batch-in-lab, and error

```
> library(MASS)
> data(coop)
> names(coop)
[1] "Lab"  "Spc"  "Bat"  "Conc"
> summary(coop)
 Lab     Spc     Bat          Conc
 L1:42   S1:36   B1:84   Min.   :0.1100
 L2:42   S2:36   B2:84   1st Qu.:0.4675
 L3:42   S3:36   B3:84   Median :1.0600
 L4:42   S4:36           Mean   :1.9215
 L5:42   S5:36           3rd Qu.:1.7000
 L6:42   S6:36           Max.   :9.9000
         S7:36
> coop2 <- coop[coop$Spc=="S1",]
> summary(coop2)
 Lab    Spc     Bat          Conc
 L1:6   S1:36   B1:12   Min.   :0.2900
 L2:6   S2: 0   B2:12   1st Qu.:0.3575
 L3:6   S3: 0   B3:12   Median :0.4000
 L4:6   S4: 0           Mean   :0.5081
 L5:6   S5: 0           3rd Qu.:0.4600
 L6:6   S6: 0           Max.   :1.3000
        S7: 0
```

Expected Mean Squares

ℓ laboratories
b batches per laboratory
r replicates per batch $n = \ell br$

$$
\begin{aligned}
E(MS(\text{ Lab})) &= br\sigma_L^2 + r\sigma_B^2 + \sigma_\epsilon^2 \\
E(MS(\text{ Batch in Lab})) &= r\sigma_B^2 + \sigma_\epsilon^2 \\
E(MS(\text{Replicate in Batch})) &= \sigma_\epsilon^2 \\
\hat{\sigma}_L^2 &= (MS_L - MS_B)/br \\
\hat{\sigma}_B^2 &= (MS_B - MSE)/r
\end{aligned}
$$

Hypothesis tests by MSL/MSB and MSB/MSE.

Analysis using lm

```
> anova(lm(Conc ~ Lab + Lab:Bat,data=coop2))
Analysis of Variance Table

Response: Conc
          Df  Sum Sq Mean Sq F value    Pr(>F)
Lab        5 1.89021 0.37804 60.0333 1.354e-10 ***
Lab:Bat   12 0.20440 0.01703  2.7049   0.02768 *
Residuals 18 0.11335 0.00630
```

The test for batch-in-lab is correct, but the test for lab is not.
The denominator should be the Lab:Bat MS, so

F(5,12) = 0.37804/0.01703 = 22.198 and p = 3.47e-4, still significant

```
Residual  0.00630  0.0794
Batch     0.00537  0.0733
Lab       0.06017  0.2453
```

We get Batch nested in Lab by including Lab:Bat without the main effect of Bat

Analysis using lmer

```
> library(lme4)

#Model below includes a fixed intercept, a random intercept per lab,
#   and a random intercept per batch

> lmer(Conc ~ 1+(1|Lab)+(1|Bat:Lab),data=coop2)
Linear mixed model fit by REML ['lmerMod']
Formula: Conc ~ 1 + (1 | Lab) + (1 | Bat:Lab)
   Data: coop2
REML criterion at convergence: -42.0432
Random effects:
 Groups   Name        Std.Dev.
 Bat:Lab  (Intercept) 0.07327
 Lab      (Intercept) 0.24529
 Residual             0.07936
Number of obs: 36, groups:  Bat:Lab, 18; Lab, 6
Fixed Effects:
(Intercept)
     0.5081
```

Hypothesis Tests

- When data are balanced, one can compute expected mean squares, and many times can compute a valid F test.
- In more complex cases, or when data are unbalanced, this is more difficult, though PROC GLM can compute expected mean squares
- One requirement for certain hypothesis tests to be valid is that the null hypothesis value is not on the edge of the possible values
- For $H_0 : \alpha = 0$, we have that α could be either positive or negative For $H_0 : \sigma^2 = 0$, negative variances are not possible

```
Effect    Variance     SD
-----------------------------------
Residual  0.00630 0.0794
Batch     0.00537 0.0733
Lab       0.06017 0.2453
```

- The variance among replicates a month apart ($0.00630 + 0.00537 = 0.01167$) is about twice that of those on the same day (0.00630), and the standard deviations are 0.1080 and 0.0794. These are CV's on the average of 21% and 16% respectively
- The variance among values from different labs is about $0.00630+0.00537+0.06017 = 0.07184$, with a standard deviation of 0.2680 and a CV of about 52%
- We would not usually conduct a formal test of whether one of the variances is 0.

Insulin Repeated Measures Example

- This is an experiment using 18 diabetic subjects.
- Nine received a protein drink and nine a non-active placebo.
- All were then challenged with a carbohydrate-heavy drink.
- Insulin levels were tracked at 0, 30, 60, and 120 minutes.

```
insulin.lmer <- lmer(log(insulin)~treat*time+(1|Subj),data=insulin)

> drop1(insulin.lmer,test="Chisq")
Single term deletions

Model:
log(insulin) ~ treat * time + (1 | Subj)
           Df    AIC    LRT   Pr(Chi)
<none>        126.91
treat:time  3 152.54 31.632 6.257e-07 ***
```

We use the log of the insulin value.
We have a random insulin level for each subject.
There is a possible shift up or down from the treatment.
There is a possible time course pattern for the two hours.
The treatment may affect times differently.
In particular, it cannot affect the time 0 level.

```
> summary(insulin.lmer)

Random effects:
 Groups   Name        Variance Std.Dev.
 Subj     (Intercept) 0.3928   0.6267
 Residual             0.1604   0.4005
Number of obs: 72, groups:  Subj, 18

Fixed effects:
                   Estimate Std. Error t value
(Intercept)         2.404553   0.247922   9.699
treatWhey          -0.050296   0.350615  -0.143
time30             -0.006161   0.188799  -0.033
time60             -0.101028   0.188799  -0.535
time120            -0.306909   0.188799  -1.626
treatWhey:time30    1.197814   0.267002   4.486
treatWhey:time60    1.394269   0.267002   5.222
treatWhey:time120   0.318168   0.267002   1.192
```

The statistically significant effects are an elevation at times 30 and 60 of the insulin levels of the treatment subjects compared to the placebo subjects.