Lecture 4. Random Effects in Completely Randomized Design

Montgomery: 3.9, 13.1 and 13.7

Random Effects vs Fixed Effects

- Consider factor with numerous possible levels
- Want to draw inference **on population of levels**
- Not just concerned with levels in experiment
- Example of differences
  - **Fixed:** Compare reading ability of 10 2nd grade classes in NY
    - Select $a = 10$ specific classes of interest.
    - Randomly choose n students from each classroom.
    - Want to compare τ_i (class-specific effects).
  - **Random:** Compare variability **among all** 2nd grade classes in NY
    - **Randomly choose** $a = 10$ classes from large number of classes.
    - Randomly choose n students from each classroom.
    - Want to assess σ^2_τ (class to class variability).
- Inference broader in random effects case: inference on population with randomly chosen levels

Random Effects Model (CRD)

- Same model as in fixed effects case

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad \begin{cases} i = 1, 2 \ldots a \\ j = 1, 2, \ldots n_i \end{cases}$$

  μ - grand mean

  τ_i - ith treatment effect

  $\epsilon_{ij} \overset{iid}{\sim} \mathrm{N}(0, \sigma^2)$

  But view treatment effects in different way

- Instead of $\sum \tau_i = 0$, assume

  $\tau_i \overset{iid}{\sim} \mathrm{N}(0, \sigma_\tau^2)$

  $\{\tau_i\}$ and $\{\epsilon_{ij}\}$ independent

- Var(y_{ij}) = $\sigma_\tau^2 + \sigma^2$

Random Effects Model

- The hypotheses are:

$$H_0: \qquad \sigma_\tau^2 = 0$$
$$H_1: \qquad \sigma_\tau^2 > 0$$

- Partitioning of Total Sum of Squares identical

  E(MS$_{\mathrm{E}}$)=σ^2

  E(MS$_{\mathrm{Treatment}}$)=$\sigma^2 + n\sigma_\tau^2$

- Under H_0, $F_0 = \mathrm{MS}_{\mathrm{Treatment}}/\mathrm{MS}_{\mathrm{E}} \sim F_{a-1,N-a}$
- Same test as before: Direct comparison of variabilities (between vs within)
- Conclusions, however, pertain to entire population

Model Estimates

- Usually interested in estimating variances
- Use mean squares (known as ANOVA method)

  $\hat{\sigma}^2 = \mathrm{MS_E}$

  $\hat{\sigma}_\tau^2 = (\mathrm{MS_{Treatment}} - \mathrm{MS_E})/n$

  If unbalanced, replace n with

  $$n_0 = ((\sum n_i)^2 - \sum n_i^2)/((a-1)\sum n_i)$$

- Estimate of σ_τ^2 can be negative
  - Supports H_0? Use zero as estimate?
  - Is the model reasonable?
  - Bayesian approach (nonnegative prior)
  - Use estimation method that only gives nonnegative estimates

Confidence intervals

- σ^2: Same as fixed case

$$\frac{(N-a)\mathrm{MS_E}}{\sigma^2} \sim \chi^2_{N-a}$$

$$\frac{(N-a)\mathrm{MS_E}}{\chi^2_{\alpha/2,N-a}} \leq \sigma^2 \leq \frac{(N-a)\mathrm{MS_E}}{\chi^2_{1-\alpha/2,N-a}}$$

- σ^2_τ: Approximate Confidence Interval

$$\hat{\sigma}^2_\tau = (\mathrm{MS_{Trt}} - \mathrm{MS_E})/n$$

  - No exact calculation of CI available.
  - Approximate CI based on Satterthwaite's Approximation:

$$r\hat{\sigma}^2_\tau / \chi^2_{\alpha/2,r} \leq \sigma^2_\tau \leq r\hat{\sigma}^2_\tau / \chi^2_{1-\alpha/2,r}$$

$$r = \frac{(\mathrm{MS_{Trt}} - \mathrm{MS_E})^2}{\mathrm{MS^2_{Trt}}/(a-1) + \mathrm{MS^2_E}/(N-a)}$$

Approximate Confidence interval

- CI for a variance component: $\sigma_0^2 = E[MS' - MS'']/k$
  - $MS' = MS_r + \cdots + MS_s$
  - $MS'' = MS_u + \cdots + MS_v$
  - No common mean squares terms shared by MS' and MS''.
  - Note $f_i MS_i/\sigma_i^2 = SS_i/\sigma_i^2 \overset{ind}{\sim} \chi^2_{f_i}$
  - Point estimate of σ_0^2: $\hat{\sigma}_0^2 = (MS' - MS'')/k$
- Satterthwaite's Approximation: $r\hat{\sigma}_0^2/\sigma_0^2 \sim \chi^2_r$

$$r = (\hat{\sigma}_0^2)^2 \Bigg/ \sum_i \frac{MS_i^2}{k^2 f_i}$$

- Approximate $(1-\alpha) \times 100\%$ CI of σ_0^2

$$r\hat{\sigma}_0^2/\chi^2_{\alpha/2,r} \le \sigma_0^2 \le r\hat{\sigma}_0^2/\chi^2_{1-\alpha/2,r}$$

- Proportion of σ_τ^2 in Var(y_{ij}), i.e., Intraclass Correlation Coefficient (ICC)

  Common estimate if goal is to reduce variance

  Uses ratio of two χ^2 distributions (i.e., F dist)

$$\frac{\mathrm{L}}{\mathrm{L}+1} \leq \frac{\sigma_\tau^2}{\sigma^2+\sigma_\tau^2} \leq \frac{\mathrm{U}}{\mathrm{U}+1}$$

$$\begin{aligned}
\mathrm{L} &= \frac{1}{n}\left(\frac{\mathrm{MS_{Trt}}}{\mathrm{MS_E} \quad F_{\alpha/2,a-1,N-a}} - 1\right) = \frac{1}{n}\left(\frac{F_0}{F_{\alpha/2,a-1,N-a}} - 1\right) \\
\mathrm{U} &= \frac{1}{n}\left(\frac{\mathrm{MS_{Trt}}}{\mathrm{MS_E} \quad F_{1-\alpha/2,a-1,N-a}} - 1\right) = \frac{1}{n}\left(\frac{F_0}{F_{1-\alpha/2,a-1,N-a}} - 1\right)
\end{aligned}$$

- Grand mean: μ

  Example: Average reading ability of 2nd grade class,

  $\overline{y}_{..} = (\overline{y}_{1.} + \overline{y}_{2.} + ... + \overline{y}_{a.})/a$.

  $\overline{y}_{i.}$ iid Normal. But what is the variance?

  CI for μ : $\overline{y}_{..} \pm t_{\alpha/2,a-1}\sqrt{\mathrm{MS_{Trt}}/(an)}$

Example

A supplier delivers several hundred batches of raw material to a company each year. The company is interested in a high yield from each batch of raw material (percent usable). Therefore, to investigate the consistency of this supplier, an experiment is done where five batches were selected at random and three yield determinations were made on each batch.

Batch				
1	2	3	4	5
74	68	75	72	79
76	71	77	74	81
75	72	77	73	79

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Between	147.73	4	36.93	20.5
Within	18.00	10	1.80	
Total	165.73	14		

- Highly significant result ($F_{.05,4,10} = 3.48$)
- $\hat{\sigma}_\tau^2 = (36.93 - 1.80)/3 = 11.71$
- 86.7% (=11.71/(11.71+1.80)) is attributable to batch differences
- Time to improve consistency of the batches

- 95% CI for σ^2

$$\left(\frac{\mathrm{SS_E}}{\chi^2_{.025,10}}, \frac{\mathrm{SS_E}}{\chi^2_{.975,10}}\right) = (18.00/20.48, 18.00/3.25) = (0.879, 5.538)$$

- 95% CI for σ^2_τ

$$\begin{aligned}
r &= \frac{(36.93 - 1.80)^2}{36.93^2/4 + 1.80^2/10} = 3.62 \\
\chi^2_{0.025,3.62} &= (1 - .62)\chi^2_{.025,3} + .62\chi^2_{.0.025,4} \\
&= .38 * 9.35 + .62 * 11.14 = 10.4598 \\
\chi^2_{0.975,3.62} &= (1 - .62)\chi^2_{.975,3} + .62\chi^2_{.0.975,4} \\
&= .38 * .22 + .62 * .48 = .3812 \\
&(3.62 * 11.71/10.4598, 3.62 * 11.71/.3812) \\
&= (4.0527, 111.2020)
\end{aligned}$$

  - SAS allows noninteger degrees of freedom for χ^2 with : `CINV(p,df)`

- 95% CI for Intraclass Correlation

$$L = \frac{1}{3} \times \left(\frac{20.52}{4.47} - 1 \right) = 1.1969 \Longrightarrow \frac{L}{L+1} = 0.5448$$

$$U = \frac{1}{3} \times \left(\frac{20.52}{1/8.84} - 1 \right) = 60.1323 \Longrightarrow \frac{U}{U+1} = 0.9836$$

$$95\% \text{ CI: } (0.545, 0.984)$$

using property that

$$F_{1-\alpha/2, v_1, v_2} = 1/F_{\alpha/2, v_2, v_1}$$

Using SAS

```
data example; input batch percent @@; cards;
  1 74   1 76   1 75   2 68   2 71
  2 72   3 75   3 77   3 77   4 72
  4 74   4 73   5 79   5 81   5 79
;

proc glm data=example;
  class batch;
  model percent=batch;
  random batch;
  output out=diag r=res p=pred;

proc gplot data=diag;
  plot res*pred;

proc mixed data=example cl;
  class batch;
  model percent = ;
  random batch / vcorr; run;
```

```
Dependent Variable: PERCENT
                                  Sum of          Mean
Source                  DF       Squares        Square   F Value     Pr > F
Model                    4     147.73333      36.93333     20.52     0.0001
Error                   10      18.00000       1.80000
Corrected Total         14     165.73333

Source                  DF     Type I SS   Mean Square   F Value     Pr > F
BATCH                    4     147.73333      36.93333     20.52     0.0001

Source      Type III Expected Mean Square
BATCH       Var(Error) + 3 Var(BATCH)
```

- GLM : Uses least squares for estimation
  - Not designed for random effects....Must compute variance estimates / CIs by hand
- MIXED: Uses restricted maximum likelihood (REML)
  - Often preferred to ML because it produces unbiased estimates of covariance parameters by taking into account the loss of degrees of freedom in estimating fixed effects
  - Usually residual variance profiled out of the likelihood

The MIXED Procedure

Estimated V Correlation Matrix for batch 1

Row	Col1	Col2	Col3
1	1.0000	0.8668	0.8668
2	0.8668	1.0000	0.8668
3	0.8668	0.8668	1.0000

Covariance Parameter Estimates

Cov Parm	Estimate	Alpha	Lower	Upper
BATCH	11.71111111	0.05	4.0450	114.2090
Residual	1.80000000	0.05	0.8788	5.5436

Fit Statistics

-2 Res Log Likelihood	62.8
AIC (smaller is better)	66.8
AICC (smaller is better)	67.8
BIC (smaller is better)	66.0

Negative σ_τ^2 Estimate Example

```
data new;
  input school subj score @@;
  cards;
1 1  74.62 1 2  73.90 1 3  72.27 1 4  71.60 1 5  73.80
1 6  77.42 1 7  72.16 1 8  76.69 1 9  75.84 1 10 70.35
2 1  72.55 2 2  71.44 2 3  72.67 2 4  72.59 2 5  71.25
2 6  68.99 2 7  69.61 2 8  77.44 2 9  73.99 2 10 73.90
3 1  76.66 3 2  74.76 3 3  70.47 3 4  75.38 3 5  68.32
3 6  76.69 3 7  73.34 3 8  68.24 3 9  69.33 3 10 78.22
;

proc glm data=new;
  class school;
  model score = school;
  random school;

proc mixed data=new cl;
  class school;
  model score = ;
  random school; run;
```

```
The GLM Procedure
                              Sum of
Source            DF         Squares   Mean Square  F Value  Pr > F
Model              2      10.1115467     5.0557733     0.60  0.5557
Error             27     227.3489500     8.4203315
Corrected Total   29     237.4604967

R-Square     Coeff Var      Root MSE    score Mean
0.042582      3.966909      2.901781      73.14967

Source            DF       Type I SS   Mean Square  F Value  Pr > F
school             2     10.11154667    5.05577333     0.60  0.5557

Source            DF     Type III SS   Mean Square  F Value  Pr > F
school             2     10.11154667    5.05577333     0.60  0.5557

Source       Type III Expected Mean Square
school        Var(Error) + 10 Var(school)
```

Parameter Estimates using ANOVA method

$$\hat{\sigma}^2 = 8.42$$

$$\hat{\sigma}_\tau^2 = \frac{5.0558-8.4203}{10} = -3.3645/10 = -0.37$$

The MIXED Procedure

Covariance Parameter Estimates

Cov Parm	Estimate	Alpha	Lower	Upper
school	0	.	.	.
Residual	8.1883	0.05	5.1935	14.7977

Fit Statistics

-2 Res Log Likelihood	146.7
AIC (smaller is better)	148.7
AICC (smaller is better)	148.8
BIC (smaller is better)	147.8

Word of Caution

- In log file for PROC MIXED analysis,

```
NOTE: Convergence criteria met.
NOTE: Estimated G matrix is not positive definite.
```

- Caused by variance estimated to be zero
- Suggests possibly removing this term from the model
- Decision often an issue in more complicated models
- Can remove positivity constraint (NOBOUND)

```
proc mixed cl nobound;
  class school;
  model score = ;
  random school ;
run;
```

- Includes null model LRT to determine if it is necessary to model the covariance structure