Nested Models

Hao Zhang

Ex) Factor A: School
Factor B: Class
- If each class is unique to a specific school, then class (B) is nested within School (A).

Nested Factors

Recall that a factor is said to be nested within another factor if its levels are observed in conjunction with just one level of the second factor.

Nested factors are usually (but not always) random factors, and they are usually blocking factors. We will see more examples in split-plot designs we will talk about later.

Example (Machine Head Experiment) One engineer wanted to study the strain readings of glass cathode supports from five different machines. Each machine had 4 heads on which the glass was formed, and the engineer decided to take four samples from each head.

Factor A machine (with 5 levels)

Factor B = head (nested within A, 4 heads per machine).

Note B is nested within A. Hence there are 20 heads. If there were crossed, there would be only 4 heads.

The two-way nested model with fixed effects is

machine effects ↑ head effects ↑

$$
\begin{aligned}
Y_{ijt} &= \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijt}, \\
\epsilon_{ijt}\ i.i.d &\sim N\left(0, \sigma^2\right), \\
t = 1, \ldots, 4; \quad i &= 1, \ldots, 5; \quad j = 1, \ldots, 4.
\end{aligned}
$$

Fixed Effect →

j(i) notation for "nested"
J is B, hence j(i). B nested within A.

where $\beta_{j(i)}$ represents the effect of jth head on the ith machine, and $j(i)$ denotes the nested structure.

In this case, the heads effects can be treated as random because the heads are regularly replaced. Then mixed effects model becomes

$$
\begin{aligned}
Y_{ijt} &= \mu + \alpha_i + B_{j(i)} + \epsilon_{ijt} \\
\epsilon_{ijt}\ i.i.d &\sim N\left(0, \sigma^2\right), \quad B_{j(i)} \sim N\left(0, \sigma^2_{B(A)}\right) \\
t = 1, \ldots, 4; \quad i &= 1, \ldots, 5; \quad j = 1, \ldots, 4.
\end{aligned}
$$

Random Effect →

Analysis of Nested Fixed Effects

Least Squares Estimates

fixed effect ↑ nested within A, also with fixed effect ↑

Consider first the simplest possible fixed-effects nested model-the two-way nested model:

$$
\begin{aligned}
Y_{ijt} &= \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijt}, \\
\epsilon_{ijt}\ i.i.d &\sim N\left(0, \sigma^2\right), \\
t = 1, \ldots, r_{ij}; \quad i &= 1, \ldots, a; \quad j = 1, \ldots, b.
\end{aligned}
$$

First note that for any $i, j, Y_{ijt}, t = 1, \ldots, r_{ij}$ independently have the normal distribution with mean $\mu + \alpha_i + \beta_{j(i)}$ and variance σ^2. Therefore, $\mu + \alpha_i + \beta_{j(i)}$ can be estimated by the sample mean $\bar{Y}_{ij.}$.

If we take an average over the subscripts t and j, we find that a comparison of the levels of A averaged over the levels of B is estimable; that is, we can estimate pairwise comparisons such as

$$
\left[\alpha_i + \bar{\beta}_{.(i)}\right] - \left[\alpha_s + \bar{\beta}_{.(s)}\right],
$$

and we can estimate general contrasts such as

$$
\sum_{i=1}^{a} c_i \left[\alpha_i + \bar{\beta}_{.(i)}\right], \quad \text{with} \sum_{i=1}^{a} c_i = 0.
$$

We can also compare the effects of those levels of B that were observed in conjunction with the same level of A; that is,

$$\left[\alpha_i + \beta_{j(i)}\right] - \left[\alpha_i + \beta_{u(i)}\right] = \beta_{j(i)} - \beta_{u(i)},$$

or, in general, $\sum_{j=1}^{b} d_j \beta_{j(i)}$, with $\sum_{j=1}^{b} d_j = 0$, for any given i.

The least squares estimator of

$$\sum_{i=1}^{a} c_i \left[\alpha_i + \bar{\beta}_{\cdot(i)}\right] \text{ is } \sum_{i=1}^{a} c_i \bar{Y}_{i\cdot\cdot}$$

Estimated by sample mean

with $\Sigma c_i = 0$. The corresponding variance is $\Sigma c_i^2 \sigma^2 / r_{i\cdot\cdot}$

Similarly, the least squares estimator of $\sum_{j=1}^{b} d_j \beta_{j(i)}$ is $\sum_{j=1}^{b} d_j \bar{Y}_{ij\cdot}$ for any i with $\Sigma d_j = 0$. The corresponding variance is $\Sigma d_j^2 \sigma^2 / r_{ij\cdot}$

Estimation of σ^2

The sum of squares for error is

$$SSE = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{t=1}^{r_{ij}} \left(Y_{ijt} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_{j(i)}\right)^2$$

$$= \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{t=1}^{r_{ij}} \left(Y_{ijt} - \bar{Y}_{ij\cdot}\right)^2$$

It follows that SSE/σ^2 has a χ^2 distribution with $(n - v)$ degrees of freedom where $v = ab$. Therefore $MSE = SSE/(n - v)$ is an unbiased estimator of σ^2.

A $100(1 - \alpha)\%$ confidence bound for σ^2 from the information in the previous subsection; that is,

$$\sigma^2 \leq \frac{ssE}{\chi^2_{n-v,1-\alpha}}.$$

Confidence intervals

Confidence intervals for $\Sigma c_i \left(\alpha_i + \bar{\beta}_{\cdot(i)}\right)$ and for $\Sigma d_j \beta_{j(i)}$ may be obtained using the relevant methods from Chap. 4 together with the formulae

$$\Sigma c_i \bar{Y}_{i\cdot\cdot} \pm w \sqrt{\sum_i \left(\frac{c_i^2}{r_i}\right) msE}$$

and

$$\Sigma d_j \bar{Y}_{ij\cdot} \pm w \sqrt{\sum_j \left(\frac{d_j^2}{r_{ij}}\right) msE}$$

$w = t_{\alpha/2, n-v}$

where w depends on the method.

Hypothesis Testing

We may test the null hypothesis that the levels of B have the same effect on the response within every given level of A, that is,

✱ meaning fixed effect

Need to specify B is nested in A.

$$H_0^{B(A)} : \{\beta_{1(i)} = \beta_{2(i)} = \ldots = \beta_{b(i)}, \text{ for every } i = 1, \ldots, a\}$$

$\beta_{j(i)} = 0$

Under this null hypothesis, the model is reduced to

$$Y_{ijt} = \mu + \alpha_i + \epsilon_{ijt}.$$

Hence

$$SSE_0 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{t=1}^{r_{ij}} \left(Y_{ijt} - \bar{Y}_{i\cdot\cdot}\right)^2$$

The difference is

$$\begin{aligned} SSB(A) &= SSE_0 - SSE \\ &= \sum_i\sum_j\sum_t \left(Y_{ijt} - \bar{Y}_{i\cdot\cdot}\right)^2 - \sum_i\sum_j\sum_t \left(Y_{ijt} - \bar{Y}_{ij\cdot}\right)^2 = \sum_i\sum_j r_{ij}\left(\bar{Y}_{ij\cdot} - \bar{Y}_{i\cdot\cdot}\right)^2. \end{aligned}$$

It follows that if $H_0^{B(A)}$ is true,

$$SSB(A)/\sigma^2 \sim \chi^2_{a(b-1)}.$$

Therefore $\frac{\text{ssB}(A)/a(b-1)}{\text{ssE}/(n-ab)}$ has an F-distribution with $a(b-1)$ and $n-ab$ degrees of freedom.

At the significance level α, reject $H_0^{B(A)}$ if

$$\frac{\text{ssB}(A)/a(b-1)}{\text{ssE}/(n-ab)} > F_{a(b-1),n-ab,\alpha}$$

.

Similarly, the decision rule for testing

$$H_0^A : \left\{\alpha_i + \bar{\beta}_{\cdot(i)} \text{ all equal}\right\}$$

against the alternative hypothesis $H_A^A : \left\{H_0^A \text{ is false}\right\}$ is

$$\text{reject } H_0^A \text{ if } \frac{\text{ssA}/(a-1)}{\text{ssE}/(n-ab)} > F_{a-1,n-ab,\alpha},$$

where

$$ssA = \sum_i r_{i\cdot}(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots})^2 = \sum_i r_{i\cdot}\bar{y}_{i\cdot\cdot}^2 - n\bar{y}_{\cdots}^2.$$

Analysis of Nested Random Effects

In the machine head experiment, we could treat the heads as a sample from the population and employ a mixed effects model:

$$\begin{aligned} Y_{ijt} &= \mu + \alpha_i + B_{j(i)} + \epsilon_{ijt} \\ \epsilon_{ijt}\ i.i.d &\sim N\left(0, \sigma^2\right), \quad B_{j(i)} \sim N\left(0, \sigma^2_{B(A)}\right) \\ t &= 1, \ldots, r_{ij}; \quad i = 1, \ldots, a; \quad j = 1, \ldots, b. \end{aligned}$$

where $a = 5, b = 4, r_{ij} = 4$.

Source of variation	Deg. of freedom	Sum of squares	Mean square	Expected mean square
A	$(a-1)$	ssA	msA	$Q(\alpha_i) + r\sigma^2_{B(A)} + \sigma^2$
				$r\sigma^2_{B(A)} + \sigma^2$
				σ^2

$B(A)$	$a(b-1)$	$ssB(A)$	$msB(A)$	$r\sigma^2_{B(A)} + \sigma^2$
Error	$ab(r-1)$	ssE	msE	σ^2
Total	$abr-1$	$sstot$		

The ANOVA table is provided below. The SSEs are obtained similarly though their expected values are different than those in the fixed effects model.

Using SAS

If factor B is nested in A, it is coded as `B(A)` in SAS. I recommend `proc mixed` if the model involves random effects. Otherwise, use `proc glm`.

First let us run a simpler but less appropriate fixed effects model in which heads nested in machine have fixed effects.

```
* machine.head.sas, machine head experiment, Table 18.3, p684;
;
* Data as arranged in Table 18.3;
data mch;
  input machine @@;
  do head=1 to 4;
    do rep=1 to 4; drop rep;
      input y @@;
      output;
  end; end;
  lines;
  1   6  2  0  8   13  3  9  8    1 10  0  6    7  4  7  9
  2  10  9  7 12    2  1  1 10    4  1  7  9    0  3  4  1
  3   0  0  5  5   10 11  6  7    8  5  0  7    7  2  5  4
  4  11  0  6  4    5 10  8  3    1  8  9  4    0  8  6  5
  5   1  4  7  9    6  7  0  3    3  0  2  2    3  7  4  0
;
run;
proc print data=mch;
run;

proc glm data=mch;
class machine head;
model y=machine head(machine);
lsmeans machine head(machine);
run;
```

The SAS output is not shown here. You shall run and look at the output.

Next, we run a mixed effects model where heads have random effects. You can use both `proc mixed` and `proc glm`.

```
proc mixed data=mch;
class machine head;
model y=machine;
random head(machine);
lsmeans machine;
run;

proc glm data=mch;
class machine head;
model y=machine head(machine);
random head(machine)/test;
lsmeans machine;
run;
```

An Illustrative Example

An experiment was reported by Gutsell (Biometrics, 1951) in which the red blood cell counts in the blood of brown trout were measured. Fish were put at random into eight troughs of water. Two troughs were assigned to each of the four levels of the treatment factor "sulfamerazine" $(0, 5, 10, 15$ grams per 100 pounds of fish added to the diet per day). After 42 days, five fish were selected at random from each trough and the red blood cell count from the blood of each fish was measured in two different counting chambers, giving two measurements per fish. The observations reported in the table below when multiplied by 5000 , give the number of red blood cells per cubic millimeter of blood.

Fish	0 gm sulf.				5 gm sulf.			
	Trough 1		Trough 2		Trough 1		Trough 2	
1	213	230	166	157	296	319	310	309
2	253	231	206	185	278	258	241	270
3	195	164	245	250	345	307	272	311
4	193	203	213	181	322	372	254	237

5	191	195	198	169	248	274	266	275
Fish	10 gm sulf.				15 gm sulf.			
	Trough 1		Trough 2		Trough 1		Trough 2	
1	339	322	196	232	278	212	287	280
2	282	285	205	186	275	311	221	243
3	236	262	252	274	186	158	331	309
4	252	209	245	216	301	281	231	244
5	263	296	249	260	223	246	292	295

Source Gutsell (1951). Copyright © 1951 International Biometric Society. Reprinted with permission

There is only one treatment factor-sulfamerazine, which has 4 levels. The four levels each were randomly assigned two troughs. Hence the troughs are the experimental units. However, the measurements are made on fish, which are the observational units. What are the sources of variation that affect these measurements? Some potential ones include trough and fish. We cannot rule out the potential differences among them. We might use the following model.

$$Y_{ijkt} = \mu + \alpha_i + B_{j(i)} + C_{k(ij)} + \epsilon_{ijkt}$$

$$\epsilon_{ijkt} \sim N\left(0, \sigma^2\right), \quad B_{j(i)} \sim N\left(0, \sigma^2_{B(A)}\right), \quad C_{k(ij)} \sim N\left(0, \sigma^2_{C(AB)}\right),$$

$$i = 1, 2, 3, 4; j = 1, 2; \quad k = 1, \ldots, 5; \quad t = 1, 2$$

where α_i is the effect of the i th level of sulfamerazine in the diet, $B_{j(i)}$ is the effect of the j th randomly selected trough assigned to the i th level of sulfamerazine, and C_k is the effect of the k th randomly selected fish from the (ij) th trough, and random variables on the right hand-side of the model are assumed to be mutually independent.

SAS code is provided below.

```
* red.blood.cell.sas, red blood cell experiment, Table 18.12, p701;
;
* Data as arranged in Table 18.12;
data redcell;
  do fish=1 to 5;
    do sulf=0,5;
      do trough=1,1,2,2;
        input count @@;
        output;
  end; end; end;
  do fish=1 to 5;
    do sulf=10,15;
      do trough=1,1,2,2;
        input count @@;
        output;
  end; end; end;
  lines;
  213 230  166 157  296 319  310 309
  253 231  206 185  278 258  241 270
  195 164  245 250  345 307  272 311
  193 203  213 181  322 372  254 237
  191 195  198 169  248 274  266 275
  339 322  196 232  278 212  287 280
  282 285  205 186  275 311  221 243
  236 262  252 274  186 158  331 309
  252 209  245 216  301 281  231 244
  263 296  249 260  223 246  292 295
;
proc print;
run;

proc mixed data=redcell;
class sulf trough fish;
model count=sulf;
random trough(sulf) fish(sulf*trough);
lsmeans sulf/cl pdiff adjust=Tukey;
run;
```

The results show a p-value of 0.0565 for the test of sulfamerazine having equal affects. The 95% simultaneous confidence intervals for the pairwise comparisons of sulfamerazine indicates a significance difference between level 1 and level 2.

The `glm` procedure will yield identical results as the `mixed` procedure if the correct error term is specified.

```
proc glm data=redcell;
class sulf trough fish;
model count=sulf trough(sulf) fish(sulf*trough);
random trough(sulf) fish(sulf*trough)/test;
```

```
lsmeans sulf/cl pdiff adjust=Tukey E=trough(sulf);
run;
```

I note that the following simpler model shall yield identical results about sulfamerazine as the previous model where the fish effects were included in the model. I

$$Y_{ijkt} = \mu + \alpha_i + B_{j(i)} + \epsilon_{ijkt}$$

$$\epsilon_{ijkt} \sim N\left(0, \sigma^2\right), \quad B_{j(i)} \sim N\left(0, \sigma^2_{B(A)}\right),$$

$$i = 1, 2, 3, 4; j = 1, 2; \quad k = 1, \ldots, 5; \quad t = 1, 2$$

The following is code for `proc mixed` for this model. I encourage you to think about why the two models produce the same results about the treatment effects.

```
proc mixed data=redcell;
class sulf trough;
model count=sulf;
random trough(sulf);
lsmeans sulf/cl pdiff adjust=Tukey;
run;
```