STAT 5200 Handout #17: Nesting and Crossing Examples (Ch. 12)

Example I is a simple nested factorial design. Example II is a bit more complicated.

Example I: Factory Layout

An industrial engineer conducts an experiment to evaluate different approaches to the insertion of electronic components on printed circuit boards. He has designed 3 assembly fixtures and 2 workplace layouts. Operators are required to perform the assembly. The layouts are in different locations in the manufacturing plant and it is not possible to use the same operators in both layouts, so 4 operators are randomly selected and assigned to each layout. Each operator produces 2 component boards with each fixture at the layout to which he or she has been assigned. The treatment combinations in the design are run in a completely random order.

1. Draw the Hasse diagram for this experiment. [Identify the factors in this experiment, which ones are fixed and which ones are random, and which factors are nested and which factors are crossed.]

Model Construction:

Factor	# Levels	Fixed or Random	Crossed / Nested / none
L = Layout	2	fixed	crossed with F
F = Fixture	3	fixed	crossed with L (and O)
O = Operator (8 operators)	4	random	nested in L
E = Error [replication]	2	random	

nested in

Layout:	1				2			
Operator:	1	2	3	4	1	2	3	4

Notation: O(L)

nested term

"parent"

So DF calculated within each level of L

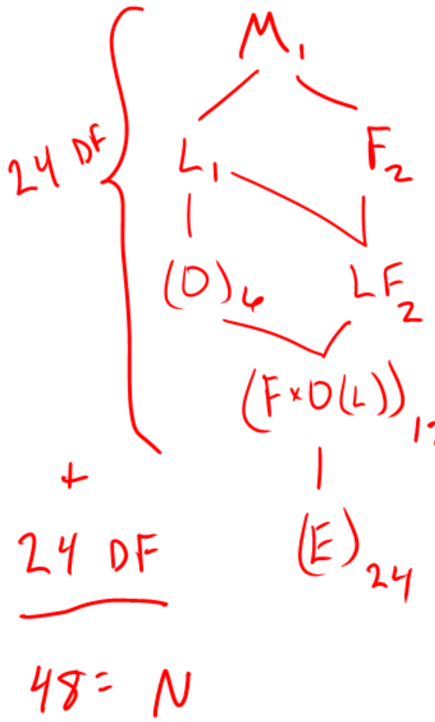
3 DF for Op. 3 DF for Op

6 DF for operator

Total sample size:

$$\underset{L}{2} \times \underset{F}{3} \times \underset{O}{4} \times \underset{rep}{2} = 48$$

↳ call this N



- No O×L interaction because:

  Operator nested in Layout (not crossed)

- F×O interaction is nested in L because:

  O is nested in L, so O is actually O(L)

- Interaction DF are product of: main effects DF

  So F×O inter. is actually F×O(L)

- Error DF are calculated after everything else above it in Hasse diagram

  N − (# DF for model, incl. M) = DF for Error

2. Write down a model for analyzing this experiment, clearly identifying all the distributional assumptions.

$$Y_{ij} = \underbrace{\mu + L_i + F_j + LF_{ij}}_{\text{Fixed}} + \underbrace{O(L)_{k(i)} + FO(L)_{jk(i)}}_{\text{Random}} + \varepsilon_{l(ijk)}$$

$i=1,2$ $\quad j=1..3$ $\quad k=1...4$ $\quad l=1,2$

$O(L)_{k(i)} \overset{iid}{\sim} N(0, \sigma^2_O)$

$FO(L)_{jk(i)} \overset{iid}{\sim} N(0, \sigma^2_{FO})$ $\qquad \varepsilon_{ijkl} \overset{iid}{\sim} N(0, \sigma^2)$

3. Compute the expected mean squares for all terms in the model.

4. Derive the *F*-ratios (and sampling distributions) for the fixed effects.

$$F_L = \frac{MS_L}{MS_O} \sim F_{1,6}$$

$$F_F = \frac{MS_F}{MS_{FO}} \sim F_{2,12}$$

$$F_{LF} = \frac{MS_{LF}}{MS_{FO}} \sim F_{2,12}$$

Recall "error" (denominator) term is next random term below in Hasse diagram (or, look at EMS ratios that equal 1 under H_o)

$E[\text{ran. var.}]$ = first moment of ran. var. → replace Expected Values with observed values

5. Construct the $\underbrace{\text{method-of-moments}}$ estimate of the $\underbrace{\text{Operator variance component}}_{}$ σ_O^2

$E[ms_O] = Ems_O$

$$Ems_O = \underbrace{\sigma^2 + 2\sigma_{FO}^2} + 6\sigma_O^2$$

want to estimate

$$Ems_{FO} = \sigma^2 + 2\sigma_{FO}^2$$

$$\text{So } (Ems_O - Ems_{FO})/6 = \sigma_O^2 \qquad \text{So } \hat{\sigma}_O^2 = \frac{ms_O - ms_{FO}}{6}$$

SAS Code: (no real data or full output here, just for reference)

```
/* STAT 5200
   simple nested factorial model
   factory layout (electronic components) data
 */




/* Read in data (here, with dummy Y values)*/
data ec; input F L O rep Y @@; cards;
    1 1 1 1 6    1 1 1 2 2    1 1 2 1 2    1 1 2 2 3
    1 1 3 1 4    1 1 3 2 2    1 1 4 1 9    1 1 4 2 3
    1 2 1 1 1    1 2 1 2 2    1 2 2 1 8    1 2 2 2 5
    1 2 3 1 4    1 2 3 2 2    1 2 4 1 7    1 2 4 2 9

    2 1 1 1 1    2 1 1 2 8    2 1 2 1 2    2 1 2 2 1
    2 1 3 1 4    2 1 3 2 2    2 1 4 1 3    2 1 4 2 4
    2 2 1 1 1    2 2 1 2 9    2 2 2 1 8    2 2 2 2 5
    2 2 3 1 7    2 2 3 2 2    2 2 4 1 2    2 2 4 2 6

    3 1 1 1 2    3 1 1 2 2    3 1 2 1 6    3 1 2 2 2
    3 1 3 1 4    3 1 3 2 2    3 1 4 1 3    3 1 4 2 4
    3 2 1 1 5    3 2 1 2 2    3 2 2 1 8    3 2 2 2 5
    3 2 3 1 1    3 2 3 2 7    3 2 4 1 2    3 2 4 2 6
    ;
run;

proc mixed data=ec method=type3;
  class F L O ;
  model y = L | F ;     ← Fixed
  random O(L) F*O(L) ;   ← Random
  title1 'Sample code, just to check EMS';
run;
```

Sample code, just to check EMS

The Mixed Procedure

Source	DF	Sum of Squares	Mean Square
L	1	22.687500	22.687500
F	2	2.000000	1.000000
F*L	2	2.000000	1.000000
O(L)	6	46.958333	7.826389
F*O(L)	12	53.666667	4.472222
Residual	24	167.500000	6.979167

no sig. evidence of a difference between Layouts

Source	Expected Mean Square	Error Term	Error DF	F Value	Pr > F
L	Var(Residual) + 2 Var(F*O(L)) + 6 Var(O(L)) + Q(L,F*L)	MS(O(L))	6	2.90	0.1395
F	Var(Residual) + 2 Var(F*O(L)) + Q(F,F*L)	MS(F*O(L))	12	0.22	0.8029
F*L	Var(Residual) + 2 Var(F*O(L)) + Q(F*L)	MS(F*O(L))	12	0.22	0.8029
O(L)	Var(Residual) + 2 Var(F*O(L)) + 6 Var(O(L))	MS(F*O(L))	12	1.75	0.1927
F*O(L)	Var(Residual) + 2 Var(F*O(L))	MS(Residual)	24	0.64	0.7874
Residual	Var(Residual)	.	.	.	.

<u>Example II: Speech Recognition</u>

A researcher wants to evaluate the hypothesis that people are better able to identify the age of a speaker who is approximately the same age as themselves. She has a large group of Speakers that are divided into 3 Agegroups (30—45, 45—60, and 60+), and by gender.

There are 5 Speakers per combination of Agegroup and Gender. She has another group of people who are Listeners, also divided into the same Agegroups and by gender. There are 6 Listeners per Agegroup and Gender combination. Each Listener listens to each Speaker exactly once and estimates the age of the Speaker. The response variable is the estimated age of the speaker minus the true age of the speaker. The researcher wants to know how this difference depends on Agegroup and Gender (of Speaker and Listener).

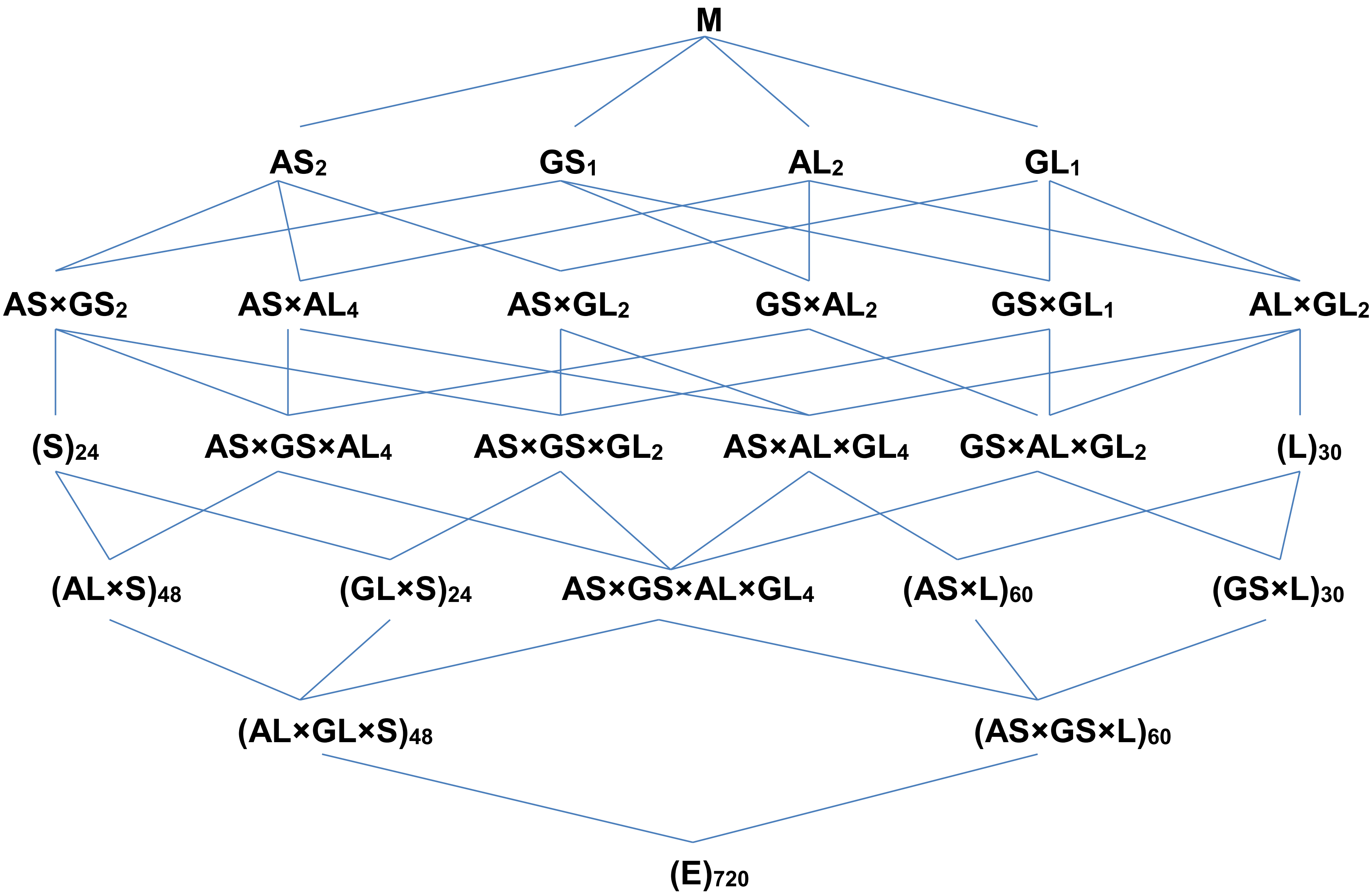
(Could do same steps 1-4 as in Example I.)

Model Construction

Factor	**# Levels**	**Fixed/Random**	**Comment**
GS = Gender of Speaker	2	Fixed	Factorial structure of these four factors
AS = Agegroup of Speaker	3	Fixed	
GL = Gender of Listener	2	Fixed	
AL = Agegroup of Listener	3	Fixed	
S = Speaker	5	Random	Nested within GS/AS combination; crossed with GL, AL, and L
L = Listener	6	Random	Nested within GL/AL combination; Crossed with GS, AS, and S

NOTES:

- DF for Speaker is (# levels Speaker – 1)×(# combinations of GS/AS crossing) because S is really S(GS×AS)
- DF for Listener is (# levels Listener – 1)×(# combinations of GL/AL crossing) because S is really L(GL×AL)
- AL×S interaction is really AL×S(GS×AS)
- GL×S interaction is really GL×S(GS×AS)
- AS×L interaction is really AS×L(GL×AL)
- GS×L interaction is really GS×L(GL×AL)
- AL×GL×S interaction is really AL×GL×S(GS×AS)
- AS×GS×L interaction is really AS×GS×L(GL×AL)
- AS×GS×AL×GL interaction term will require an approximate F-test
- Here there is only one observation per combination of S and L, so error term is interaction with fullest set of subscripts. [i.e., E = S×L(AS×GS×AL×GL) due to lack of S×L replication]


M
AS2
GS1
AL2
GL1
AS×GS2
AS×AL4
AS×GL2
GS×AL2
GS×GL1
AL×GL2
(S)24
AS×GS×AL4
AS×GS×GL2
AS×AL×GL4
GS×AL×GL2
(L)30
(AL×S)48
(GL×S)24
AS×GS×AL×GL4
(AS×L)60
(GS×L)30
(AL×GL×S)48
(AS×GS×L)60
(E)720