Lecture 26 Friday March 31

Lecture notes pp. 126 – 131

<u>TOPICS</u> Two-way Balanced ANOVA

Decomposition of sums of squares & df

Analogy(ies) w/ one-way ANOVA

Mean squares

Expected mean squares

Corn Yield Ex.

Recall (pp. 72, 73):

One-way ANOVA, decomposition of deviations, sums of squares, degrees of freedom

For $i=1,\dots,r$, $j=1,\dots,n_i$: $\quad Y_{ij} - \bar{Y}_{..} = \bar{Y}_{i.} - \bar{Y}_{..} + Y_{ij} - \bar{Y}_{i.}$

$$\sum_{i=1}^{r} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{..})^2 = \sum_{i=1}^{r} n_i (\bar{Y}_{i.} - \bar{Y}_{..})^2 + \sum_{i=1}^{r} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i.})^2$$

$$SSTO = SSTR + SSE$$

$$n_T - 1 = r - 1 + n_T - r$$

Before p. 126

Decomposition of deviations, sums of squares, and degrees of freedom
(first stage)

First, get the decompositions as if we were just doing a one-way ANOVA, where the single treatment factor is defined by all possible pairs of Factors A and B, and which has $r = ab$ levels.

$$Y_{ijk} - \overline{Y}_{\ldots} = (\overline{Y}_{ij.} - \overline{Y}_{\ldots}) + (Y_{ijk} - \overline{Y}_{ij.})$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk} - \overline{Y}_{\ldots})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b} n(\overline{Y}_{ij.} - \overline{Y}_{\ldots})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk} - \overline{Y}_{ij.})^2$$

$$\text{SSTO} = \text{SSTR} + \text{SSE}$$ (like one-way ANOVA)

$$nab - 1 = (ab - 1) + (n - 1)ab$$

$n_T - 1$ $\quad$ $r - 1$ $\quad$ $n_T - r$

In the second stage of the decomposition of SS, the treatment sum of squares is broken down into parts due to main effect of A, main effect of B, and interaction effect AB. This stage is crucial in two-way ANOVA.

Decomposition of treatment deviations, SS, df For $i = 1, \ldots, a$; $j = 1, \ldots, b$; $k = 1, \ldots, n$:

1 $$\overline{Y}_{ij.} - \overline{Y}_{...} = \underbrace{(\overline{Y}_{i..} - \overline{Y}_{...})}_{\hat{\alpha}_i} + \underbrace{(\overline{Y}_{.j.} - \overline{Y}_{...})}_{\hat{\beta}_j} + \underbrace{(\overline{Y}_{ij.} - \overline{Y}_{i..} - \overline{Y}_{.j.} + \overline{Y}_{...})}_{(\hat{\alpha\beta})_{ij}}$$

2 $$n\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{Y}_{ij.} - \overline{Y}_{...})^2 = nb\sum_{i=1}^{a}(\overline{Y}_{i..} - \overline{Y}_{...})^2 + na\sum_{j=1}^{b}(\overline{Y}_{.j.} - \overline{Y}_{...})^2 + n\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{Y}_{ij.} - \overline{Y}_{i..} - \overline{Y}_{.j.} + \overline{Y}_{...})^2$$

3 df: $$ab - 1 = (a-1) + (b-1) + (a-1)(b-1)$$

---

2 We can derive Eqn. 2 from Eqn. 1, by the fact that when you square the r.h.s. of 1 and sum over all nab observations, all the (three of) cross-product terms vanish.

We have:

$$\text{SSA} = nb\sum_{i=1}^{a}(\overline{Y}_{i..} - \overline{Y}_{...})^2,$$

which is SSTR from one-way ANOVA on Factor A ignoring Factor B

$$\text{SSB} = na\sum_{j=1}^{b}(\overline{Y}_{.j.} - \overline{Y}_{...})^2,$$

which is SSTR from one-way ANOVA on Factor B ignoring Factor A

$$\text{SSAB} = n\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{Y}_{ij.} - \overline{Y}_{i..} - \overline{Y}_{.j.} + \overline{Y}_{...})^2$$

<u>Explain why df for the AB interaction is df(AB) $= (a-1)(b-1)$</u>

By example. Take a $2x3$ design with $(\alpha\beta)_{11} = 2$, $(\alpha\beta)_{12} = 1$. Note that all remaining interaction effects are determined by the constraints on the interaction parameters. So the "degrees of freedom for interaction" is 2 here, or $(2-1)(3-1)$.

$(a-1)(b-1)$

Inter	B1	B2	B3
A1	2	1	-3
A2	-2	-1	3

Now $(\alpha\beta)_{13} = -2-1 = -3$

And $(\alpha\beta)_{21} = -(\alpha\beta)_{11} = -2$ etc

Mean Squares

Sum of squares divided by d.f.
These are statistics

$$\text{MSE} = \frac{SSE}{ab(n-1)}$$

$$\text{MSA} = \frac{SSA}{a-1}$$

$$\text{MSB} = \frac{SSB}{b-1}$$

$$\text{MSAB} = \frac{SSAB}{(a-1)(b-1)}$$

Expected Mean Squares

These are probability facts about sampling distr's.

$$E(\text{MSE}) = \sigma^2$$

$$E(\text{MSA}) = \sigma^2 + nb\frac{\sum \alpha_i^2}{a-1}$$

$$\frac{\sum_{i=1}^{a} \alpha_i^2}{a-1}$$

$$E(\text{MSB}) = \sigma^2 + na\frac{\sum \beta_j^2}{b-1}$$

$$E(\text{MSAB}) = \sigma^2 + n\frac{\sum\sum(\alpha\beta)_{ij}^2}{(a-1)(b-1)}$$

Example. Corn Yield, true cell means (p. 111) We don't know the error variance σ^2.

	Fertilizer			
Manure	Low	High	$\mu_{i.}$	α_i
Low	11.0	14.0	12.5	−1.0
High	14.0	15.0	14.5	1.0
$\mu_{.j}$	12.5	14.5	$\mu_{..} = 13.5$	
β_j	−1.0	1.0		

$$\text{E(MSE)} = \sigma^2$$

$$\text{E(MSA)} = \sigma^2 + 5(2)\frac{[(-1)^2 + 1^2]}{(2-1)} = \sigma^2 + 20$$

$$\text{E(MSB)} = \sigma^2 + 5(2)\frac{[(-1)^2 + 1^2]}{(2-1)} = \sigma^2 + 20$$

$$\text{E(MSAB)} = \sigma^2 + n \frac{\sum\sum(\alpha\beta)_{ij}^2}{(a-1)(b-1)}$$

$$= \sigma^2 + 5 \frac{((-.5)^2 + (.5)^2 + (-.5)^2 + (.5)^2)}{(1)(1)}$$

$$= \sigma^2 + 5$$

ANOVA table

Source of Variation	Sum of Squares	df	Mean Square
Factor A	SSA	$a-1$	$\frac{\text{SSA}}{a-1}$
Factor B	SSB	$b-1$	$\frac{\text{SSB}}{b-1}$
AB interaction	SSAB	$(a-1)(b-1)$	$\frac{\text{SSAB}}{(a-1)(b-1)}$
Error	SSE	$ab(n-1)$	$\frac{SSE}{ab(n-1)}$
Total	SSTO	$nab-1$	

where

$$\text{SSA} = nb \sum_{i=1}^{a} (\overline{Y}_{i..} - \overline{Y}_{...})^2$$

$$\text{SSB} = na \sum_{j=1}^{b} (\overline{Y}_{.j.} - \overline{Y}_{...})^2$$

$$\text{SSAB} = n \sum_{i=1}^{a} \sum_{j=1}^{b} (\overline{Y}_{ij.} - \overline{Y}_{i..} - \overline{Y}_{.j.} + \overline{Y}_{...})^2$$

$$\text{SSE} = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (Y_{ijk} - \overline{Y}_{ij.})^2$$

There are three F tests

Question 1 for ANOVA Analysis First question we examine in the statistical analysis is: Is there an interaction between Factor A and Factor B? That is, does the effect of Factor A depend on what is the level of Factor B?

Ex. Corn Yield

Cell Means (DATA)

	Fertilizer		(Simple) Individual Measures
Manure	Low	High	of effect
Low	11.3	13.9	$13.9 - 11.3 = 2.6$
High	14.0	15.1	$15.1 - 14.0 = 1.1$
Individual Measures of effect	$14.0 - 11.3 = 2.7$	$15.1 - 13.9 = 1.2$	

The individual measures of effect are not equal, for this data.

Question: Are the observed differences just due to chance?

We will fit the model in R; the usual output includes the F test for interaction, which is what is needed here.