Lecture 24

Wed March 22, 2023

PLAN

- ✓ Hw5 example like lightning
- ✓ Two-way factor effects model
- ✓ Data notation, SS, df
- Ex. Corn Yield

next time

EX. (like Hw5)

Factor A Gender	Factor B - Age Young	Middle	Old
Male	11	13	18
Female	7	9	14

Chart of μ_{ij}'s Learning time cell means

Q#1

a) $a=2, \quad b=3$

b) $\mu_{\cdot\cdot} = \sum_{i=1}^{2}\sum_{j=1}^{3} \mu_{ij}/6 = \dfrac{11+13+18+7+9+14}{6} = 12$

c) $\mu_{1\cdot} = \sum_{j=1}^{3} \mu_{1j}/3 = \dfrac{11+13+18}{3} = 14, \quad \mu_{2\cdot} = \sum_{j=1}^{3} \mu_{2j}/3 = 10$

$\alpha_1 = \mu_{1\cdot} - \mu_{\cdot\cdot} = 14-12 = 2, \quad \alpha_2 = \mu_{2\cdot} - \mu_{\cdot\cdot} = 10-12 = -2$

d) $\mu_{\cdot 1} = \sum_{i=1}^{2} \mu_{i1}/2 = \dfrac{11+7}{2} = 9, \quad \mu_{\cdot 2} = \sum_{i=1}^{2} \mu_{i2}/2 = 11, \quad \mu_{\cdot 3} = \sum_{i=1}^{2} \mu_{i3}/2 = 16$

$\beta_1 = \mu_{\cdot 1} - \mu_{\cdot\cdot} = 9-12 = -3, \quad \beta_2 = \mu_{\cdot 2} - \mu_{\cdot\cdot} = 11-12 = -1, \quad \beta_3 = \mu_{\cdot 3} - \mu_{\cdot\cdot} = 4$

e) $(\alpha\beta)_{11} = \mu_{11} - (\mu_{\cdot\cdot} + \alpha_1 + \beta_1) = \mu_{11} - \mu_{1\cdot} - \mu_{\cdot 1} + \mu_{\cdot\cdot} = 11-14-9+12 = 0$

$(\alpha\beta)_{12} = \mu_{12} - \mu_{1\cdot} - \mu_{\cdot 2} + \mu_{\cdot\cdot} = 13-14-11+12 = 0$

$(\alpha\beta)_{13} = \mu_{13} - \mu_{1\cdot} - \mu_{\cdot 3} + \mu_{\cdot\cdot} = 18-14-16+12 = 0$

$(\alpha\beta)_{21} =$

$(\alpha\beta)_{22} =$

$(\alpha\beta)_{23} =$

Q#2// $\mu_{12} - \mu_{11} = 2$, & $\mu_{13} - \mu_{12} = 5$: Does not imply Factors A and B interact because these are two different simple effects of Factor B at the same (Male) level of Factor A. To study interaction compare the same effect of B at different levels of A.

Recall: One-way ANOVA, factor effects model (p. 81)

$$Y_{ij} = \mu_\cdot + \tau_i + \varepsilon_{ij}, \quad i=1,\dots,r;\ j=1,\dots,n_i$$

where

- $\mu_\cdot = \dfrac{\sum_{i=1}^{r} \mu_i}{r}$

- τ_i is the effect of the i^{th} factor level
  $\tau_i = \mu_i - \mu_\cdot$

- ε_{ij} iid $N(0, \sigma^2)$

Constraint on τ_i's $\quad \sum_{i=1}^{r} \tau_i = 0$

Two-way ANOVA, factor effects model

The k^{th} observation in cell (i,j) is denoted Y_{ijk}, $i = 1, \ldots, a$, $j = 1, \ldots, b$, $k = 1, \ldots, n$, and the model for this observation says:

equal # observations per trt

$$Y_{ijk} = \underbrace{\overbrace{\boxed{\mu_{\cdot\cdot} + \alpha_i + \beta_j + (\alpha\beta)_{ij}}}^{\mu_{ij}}}_{\text{fixed, unknown parameters}} + \epsilon_{ijk}$$

→ random error

where

$\mu_{\cdot\cdot}$ is a constant,

$\sum_{i=1}^{a} \alpha_i = 0$

$\sum_{j=1}^{b} \beta_j = 0$

$\sum_{j=1}^{b} (\alpha\beta)_{ij} = 0$ for all $i = 1, \ldots, a$ For ~~all the~~ each row, interaction effects sum to 0

$\sum_{i=1}^{a} (\alpha\beta)_{ij} = 0$ for all $j = 1, \ldots, b$ For each column, " " " " "

Assumptions on the ϵ_{ijk}: independent, normally distributed, with mean 0 and constant variance σ^2

No. of Parameters in model:		# Constraints
$\mu_{\cdot\cdot}$	$-$ 1	1
α_i	$-$ a	1
β_j	$-$ b	$a+b$
$(\alpha\beta)_{ij}$	$-$ ab	

Total # $-$ $1 + a + b + ab$

We can estimate <u>ab</u> parameters. So, we have $a+b+1$ too many parameters

Need $a+b+1$ <u>constraints</u>.

We have $a+b+2$ constraints

Notation for Data

Let $Y_{ij.} = \sum_{k=1}^{n} Y_{ijk}$, be the sum of the observations for the treatment corresponding to the i^{th} level of factor A and the j^{th} level of factor B.

The corresponding mean is:

$$\overline{Y}_{ij.} = \frac{1}{n} Y_{ij.}$$

The total of all observations for the i^{th} level of factor A is:
$Y_{i..} = \sum_{j=1}^{b} \sum_{k=1}^{n} Y_{ijk}$

and the corresponding mean is:

$$\overline{Y}_{i..} = \frac{1}{bn} Y_{i..}$$

Also $\overline{Y}_{i..} = \frac{1}{b} \sum_{j=1}^{b} \overline{Y}_{ij.}$ because :

reason $\overline{Y}_{i..} = \frac{1}{bn} Y_{i..} = \frac{1}{b} \frac{1}{n} \sum_{j=1}^{b} Y_{ij.} = \frac{1}{b} \sum_{j=1}^{b} \frac{1}{n} Y_{ij.}$

$= \frac{1}{b} \sum_{j=1}^{b} \overline{Y}_{ij.}$

Notation, con.

The total of all observations for the j^{th} level of factor B is:
$Y_{.j.} = \sum_{i=1}^{a} \sum_{k=1}^{n} Y_{ijk}$
and the corresponding mean is:

$$\overline{Y}_{.j.} = \frac{1}{an} Y_{.j.}$$

The sum of all observations in the study is:

$$Y_{\dots} = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} Y_{ijk}$$

and the overall mean is

$$\overline{Y}_{\dots} = \frac{1}{abn} Y_{\dots}$$

<u>Show that $\overline{Y}_{\dots}$</u> = average of the row averages, $\overline{Y}_{i..}$ $i=1,\dots,a$

Note that $\overline{Y}_{\dots} = \frac{1}{abn} \sum_{i=1}^{a} Y_{i..} = \frac{1}{a} \sum_{i=1}^{a} \frac{Y_{i..}}{bn} = \frac{1}{a} \sum_{i=1}^{a} \overline{Y}_{i..}$