STA 303/1002

Note Title 1/26/2012

Two-Way Classification
& Two-Way Analysis of Variance

Continuous Response, 2 factors

Example: The Pygmalion Effect

Pygmalion effect - high expectations of a supervisor translate to improved performance by subordinate

Experiment to test Pygmalion effect

10 companies of soldiers in army training camp
Each company has 3 platoons ; each platoon trains together under one leader (one leader per platoon)

Within each company, 1 platoon leader was told he had an exceptionally good group. — this is the pygmalion platoon ; the other 2 within each company are controls platoons
(which platoon chosen to be pygmalion was chosen randomly)

Response variable: score on a basic weapons test per platoon

Factors (explanatory variables)

company - 10 levels

treatment - 2 levels (Pygmalion, control)

Looking at interaction plot (produced with proc sgplot)

- mean of each (factor) treatment combination (20 means)

Looks like difference between control and pygmalion groups differs with company

- so we want to investigate the interaction between company and treatment.

We'll fit a Linear Model:

$Y_i$ - score for $i^{th}$ platoon

Explanatory variables: Indicator variables:

9 for company (since 10 levels)

1 for treatment (since 2 levels)

9 interaction terms ($9 = 9 \times 1$)

Define $I_{PYG,i} = \begin{cases} 1 & \text{if } i^{th} \text{ observation is Pygmalion} \\ 0 & \text{if control} \end{cases}$

$I_{COMPk,i} = \begin{cases} 1 & \text{if } i^{th} \text{ observation is company } k \\ 0 & \text{otherwise} \end{cases}$

$N = 29$ ( company 3 only has 2 platoons)

<u>Model</u>: 

$$
\begin{aligned}
Y_i = \beta_0 &+ \beta_1 I_{PYG,i} \\
&+ \beta_2 I_{comp1,i} &&+ \beta_{11} I_{PYG,i} * I_{comp1,i} \\
&+ \beta_3 I_{comp2,i} &&+ \beta_{12} I_{PYG,i} * I_{comp2,i} \\
&+ \cdots &&+ \cdots \\
&+ \beta_{10} I_{comp9,i} &&+ \beta_{19} I_{PYG,i} * I_{comp9,i} \\
& &&+ e_i
\end{aligned}
$$

Mean of $Y_i$ for each factor combination:
$E(Y_i)$

| Company | Pygmalion | Control | Treatment effect (Pyg − Control) |
|---|---|---|---|
| 1 | $\beta_0 + \beta_1 + \beta_2 + \beta_{11}$ | $\beta_0 + \beta_2$ | $\beta_1 + \beta_{11}$ |
| 2 | $\beta_0 + \beta_1 + \beta_3 + \beta_{12}$ | $\beta_0 + \beta_3$ | $\beta_1 + \beta_{12}$ |
| 3 | $\beta_0 + \beta_1 + \beta_4 + \beta_{13}$ | $\beta_0 + \beta_4$ | $\beta_1 + \beta_{13}$ |
| 4 | ⋮ | ⋮ | ⋮ |
| 5 | | | |
| 6 | | | |
| 7 | | | |
| 8 | | | |
| 9 | | | |
| 10 | $\beta_0 + \beta_1$ | $\beta_0$ | $\beta_1$ |

Question 1 Does mean treatment effect differ with company?

$H_0: \beta_{11} = \beta_{12} = \dots = \beta_{19} = 0$

vs $H_a$: at least one not-zero

Partial F-test In linear model, test whether subset of coefficients of explanatory variables are 0.

Approach: Fit full and reduced model

full → all explanatory variables

reduced → without variables whose coefficients you are testing

Test statistic: $F_{obs} = \dfrac{\dfrac{SSReg_{full} - SSReg_{reduced}}{\#\text{ of } \beta\text{'s being tested}}}{MSE_{full}}$

If $H_0$ is true, $F_{obs}$ is an observation from an $F$ distribution with df = # of $\beta$'s being tested, df of error for full model

SAS: proc glm
model score = company|treatment;

equivalent to

```
model score = company treatment
              company * treatment;
```