Lecture 16 Friday February 17, 2023

Notes pp. 85 - 89

Ch. 17 Analysis of Factor Level Means

- Individual inferences
- Contrasts (an example)
- Kenton Food Company Example

Chapter 17 Analysis of Factor Level Means

It is convenient to use the cell-means formulation of the model for this discussion.

Topics:

- Individual inferences, e.g. for a single μ_i or for a difference of means $\mu_i - \mu_j$. → Not "multiple comparisons," not "simultaneous inference"
- Contrasts → particular type of linear combination of cell means
- Simultaneous inference → Needed for usual follow-up analysis after a significant F test in order to "protect" Type I error for many tests at once

Individual inferences

Ex: Kenton Food Company (See p. 59)

Package Design	Characteristics
1	3 colors, with cartoons
2	3 colors, without cartoons
3	5 colors, with cartoons
4	5 colors, without cartoons

CRD
one factor
at
4 levels

factorial design, two factors
at two levels each

Kenton Food Company We got the following summary statistics (p. 71, p. 76):

Package Design	Number of Stores	Mean	
i	n_i	$\bar{Y}_{i.}$	mean # cases sold
1	5	14.6	
2	5	13.4	
3	4	19.5	
4	5	27.2	

MSE $= \hat{\sigma}^2 = 10.55$

(Note: We find the value of MSE from the ANOVA table, row labelled Error or Residuals, Mean Square column.)

df(Error) $= n_T - r = 19 - 4 = 15$

Confidence Interval for a factor level mean μ_i

The goal is as in the one-sample situation, but since we have data from several groups, we use all the data to estimate the variance σ^2, with $\hat{\sigma}^2 = \text{MSE} = \text{SSE}/(n_T - r)$.

Basic Facts

1. $E(\hat{\mu}_i) = \mu_i$. $\hat{\mu}_i = \bar{Y}_{i\cdot}$
2. $\text{SD}(\hat{\mu}_i) = \dfrac{\sigma}{\sqrt{n_i}}$
3. $\hat{\mu}_i$ has a normal distribution.
4. $(\hat{\mu}_i - \mu_i)\Big/\left(\dfrac{\sigma}{\sqrt{n_i}}\right) \sim \mathcal{N}(0,1)$.
5. $(\hat{\mu}_i - \mu_i)\Big/\left(\dfrac{\hat{\sigma}}{\sqrt{n_i}}\right) \sim t_{n_T - r}$. $\hat{\sigma} = \sqrt{MSE}$

See p. 15, Useful Facts about linear combinations of normal r.v.'s.

See p. 28 for **Basic Facts** about the probability background in the two-sample scenario for comparison of two means. The same structure is used for the outline of facts in the current situation.

KEY POINT: Use all the data, and all the degrees of freedom for error in the one-way ANOVA, for carrying out inference about the i^{th} group mean.

It follows from point 5 that a $100(1-\alpha)\%$ CI for μ_i is:

Ex. Kenton Food Company

- Estimate the mean sales for package design 4 with a 95% CI:

- Estimate the mean difference in sales for package design 2 vs. package design 1.

It follows from point 5 that a $100(1-\alpha)\%$ CI for μ_i is:
$\hat{\mu}_i \pm t_{n_T-r,1-\frac{\alpha}{2}} \times \mathrm{SE}(\hat{\mu}_i)$, where $\hat{\mu}_i = \bar{Y}_{i.}$, $\mathrm{SE}(\hat{\mu}_i) = \hat{\sigma}/\sqrt{n_i}$.

Ex. Kenton Food Company
Estimate the mean sales for package design 4 with a 95% CI:
A 95% CI for μ_4 is (24.1, 30.3)

First we write down the ingredients: $\bar{Y}_{4.} = 27.2$, $t_{15,.975} = 2.131$, MSE $= 10.547$, $n_4 = 5$

Now obtain the CI:

$27.2 \pm 2.131 \times \sqrt{\frac{10.547}{5}}$,
$27.2 \pm 2.131(1.4524)$
27.2 ± 3.0951

margin of error

Ex. Kenton Food Company

Estimate the mean difference in sales for package design ~~2~~ 1 vs. package design ~~1~~ 2.

Method: An individual $100(1-\alpha)\%$ CI for $\mu_1 - \mu_2$ is $\hat{\mu}_1 - \hat{\mu}_2 \pm t_{n_T - r, 1-\frac{\alpha}{2}} \times$ SE$(\hat{\mu}_1 - \hat{\mu}_2)$

Ingredients:

$$\hat{\mu}_1 - \hat{\mu}_2 = \bar{Y}_{1\cdot} - \bar{Y}_{2\cdot} = 14.6 - 13.4 = 1.2$$

$$SE(\hat{\mu}_1 - \hat{\mu}_2) = \hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = \sqrt{10.547}\sqrt{\frac{1}{5} + \frac{1}{5}} = 2.054$$

The 95% CI for $\mu_1 - \mu_2$:

$1.2 \pm 2.131(2.054)$, or

1.2 ± 4.378, which is

$(-3.18, 5.58)$

Kenton, CI for $\mu_1 - \mu_2$:

Interpretation of the 95% CI: We are 95% confident that $\mu_1 - \mu_2$ is between -3.2 and 5.6. We fail to reject $H_0 : \mu_1 - \mu_2 = 0$ at level $\alpha = .05$, because zero is in the 95% CI. The value zero is a "plausible value" for the difference between these two population mean sales volumes, for Package Designs 1 and 2.

- Let's compare mean sales for designs with and without cartoons; estimate the following contrast.

parameter → $$L = \frac{\mu_1 + \mu_3}{2} - \frac{\mu_2 + \mu_4}{2}$$

~~ave~~ w/ cartoons ignoring color by averaging over the groups w/ 3 & 5 colors

w/out cartoons ignoring color

For now, be aware that a *contrast* of cell means is a particular type of linear combination of the set of cell means. The set is $\{\mu_1, \mu_2, \dots, \mu_r\}$

Let's state the method for forming a $100(1-\alpha)\%$ CI for *any* linear combination of the cell means.

Setting: One-way ANOVA, r groups, population means $\mu_1, \mu_2, \dots, \mu_r$, error variance σ^2, MSE denoted $\hat{\sigma}^2$, group sample sizes $n_1, n_2, \dots, n_r$, r constants $c_1, c_2, \dots, c_r$.

Let $L = \sum_{i=1}^{r} c_i \mu_i$ be a linear combination of the cell means.

BASIC FACTS

let $\hat{L} = \sum_{i=1}^{r} c_i \bar{Y}_{i\cdot}$

1. $E(\hat{L}) = L$
2. $SD(\hat{L}) = \sigma \sqrt{\sum_{i=1} \frac{c_i^2}{n_i}}$ ← Derive this
3. ~~not~~ $\hat{L} \sim$ normal
4. $\frac{\hat{L} - L}{SD(\hat{L})} \sim N(0,1)$
5. $\frac{\hat{L} - L}{\hat{\sigma} \sqrt{\sum^{r} c_i^2 / n_i}} \sim t_{n_T - r}$

From Basic Fact 5

A $100(1-\alpha)\%$ CI for L is:

$$\hat{L} \pm t_{n_T - r,\, 1-\frac{\alpha}{2}} \times SD(\hat{L})$$

A convenient way to refer to ~~SD~~ variability of $\hat{L}$:

$$Var(\hat{L}) = \sigma^2 \sum_{i=1}^{r} \frac{c_i^2}{n_i}$$ true variance of $\hat{L}$, a parameter

and

$$\hat{Var}(\hat{L}) = MSE \sum_{i=1}^{r} \frac{c_i^2}{n_i}$$ estimated variance of $\hat{L}$, a statistic

And we use the notation

$$\hat{Var}(\hat{L}) = s^2(\hat{L}), \qquad s(\hat{L}) = \sqrt{s^2(\hat{L})}$$

$$L = \frac{\mu_1 + \mu_3}{2} - \frac{\mu_2 + \mu_4}{2}$$

Cell means	μ_1,	μ_2,	μ_3,	μ_4
c_i	$\frac{1}{2}$	$\frac{-1}{2}$	$\frac{1}{2}$	$\frac{-1}{2}$
n_i	5	5	4	5

$$Var(\hat{L}) = \sigma^2 \left(\left(\frac{1}{2}\right)^2 \frac{1}{5} + \left(\frac{-1}{2}\right)^2 \frac{1}{5} + \left(\frac{1}{2}\right)^2 \frac{1}{4} + \left(\frac{-1}{2}\right)^2 \frac{1}{5} \right)$$

$$= \sigma^2 \left(\frac{1}{20} + \frac{1}{20} + \frac{1}{16} + \frac{1}{20} \right)$$

$$= \sigma^2 \frac{17}{80}$$

$$\hat{Var}(\hat{L}) = MSE \left(\frac{17}{80} \right)$$

89-3