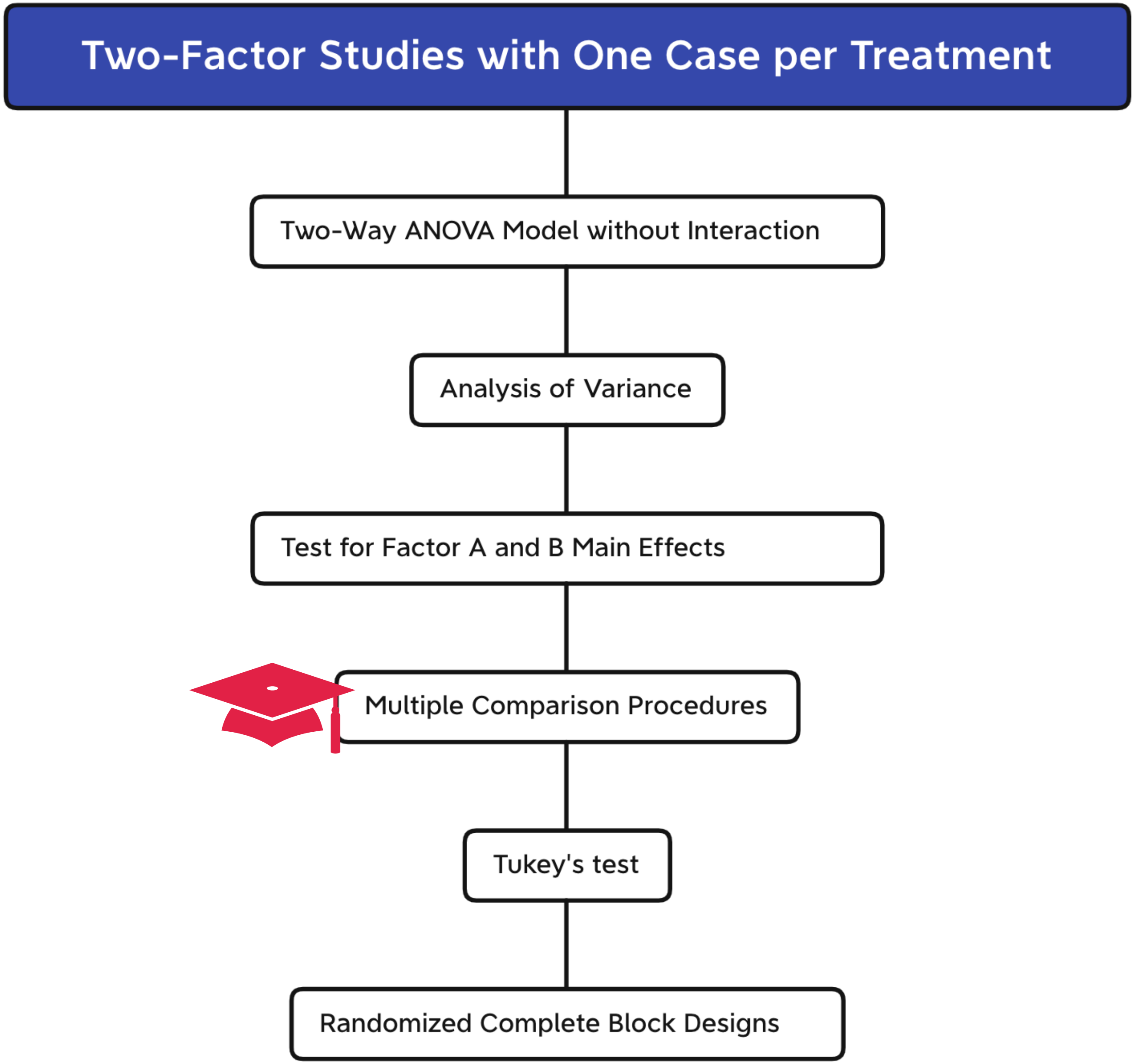
Lecture 5: Two-Factor Studies with One Case per Treatment

STA 106: Analysis of Variance

Xiner Zhou
Department of Statistics, University of California, Davis


Two-Factor Studies with One Case per Treatment
Two-Way ANOVA Model without Interaction
Analysis of Variance
Test for Factor A and B Main Effects
Multiple Comparison Procedures
Tukey's test
Randomized Complete Block Designs

Analysis of Factor A (and B) Main Effects (When Factors Do Not Interact)

Multiple Comparison Procedure: Bonferroni

Suppose we're interested in making infernece about multiple quantities,

that are linear combinations of factor A level means (or factor B level means),

i.e., a family containing g linear combinations of factor level means

$$\mathscr{L} = \{L_1 = \sum_{i=1}^{r} c_{1i}\mu_{i\cdot}, \ldots, L_g = \sum_{i=1}^{r} c_{gi}\mu_{i\cdot}\}$$

$$\hat{L} = \sum_i c_i \bar{Y}_{i\cdot} \qquad s^2(\hat{L}) = \frac{MSE}{b}\sum_i c_i^2$$

Bonferroni's idea:
One very easy and conservative way to control family-wise error rate at α is to control individual test's significance level at $\alpha_0 = \dfrac{\alpha}{g}$

This procedure includes any inference about a single quantity as special case, just take g=1.

Analysis of Factor A (and B) Main Effects (When Factors Do Not Interact)

Multiple Comparison Procedure: Bonferroni

$(1-\alpha)100\,\%$ confidence interval for individual quantity in this family:

$$\hat{L}_i \pm Bs\left(\hat{L}_i\right) \text{ for } i = 1 \ldots g$$

$$B = t\left(1 - \frac{\alpha}{2g}; (a-1)(b-1)\right)$$

Guarantee:
family-wise confidence coefficient is at least $(1-\alpha)100\,\%$

Meaning:
in at least $(1-\alpha)100\,\%$ of repetition of experiments, all the intervals in the family cover the true corresponding $L_i's$ $\alpha\ \%$ of repetition of experiments, some tests in the family made false discovery when the null hypothesis was true.

Hypothesis testing (t-test) for individual quantity in this family:

$$H_0^i : L_i = 0 \ H_a^i : L_i \neq 0$$

$$t^* = \frac{\hat{L}_i}{s(\hat{L}_i)} \sim t_{n_T - r} \text{if } H_0 \text{ is true}$$

If $|t^*| \leq B$, conclude H_0

If $|t^*| > B$, conclude H_a

Guarantee:
family-wise Type I error is at most α

Meaning:
in at most $\alpha\ \%$ of repetition of experiments, some tests in the family made false discovery when the null hypothesis was true.

Analysis of Factor A (and B) Main Effects (When Factors Do Not Interact)

Multiple Comparison Procedure: Sheffe

Suppose we're interested in making inference about <u>all possible contracts of factor A level means</u>
i.e., a family containing all possible contracts of factor A level means

$$\mathscr{L} = \{L = \sum_{i=1}^{r} c_i \mu_{i\cdot} \text{ where } \sum_{i=1}^{r} c_i = 0\}$$

Infinitely many claims or quantities

$$\hat{L} = \sum_i c_i \bar{Y}_{i\cdot\cdot} \qquad s^2(\hat{L}) = \frac{MSE}{b} \sum_i c_i^2$$

Analysis of Factor A (and B) Main Effects (When Factors Do Not Interact)

Multiple Comparison Procedure: Sheffe

$(1-\alpha)100\ \%$ confidence interval for individual quantity in this family:

$$\hat{L}_i \pm Ss\left(\hat{L}_i\right)$$

$$S = \sqrt{(a-1)F(1-\alpha; a-1, (a-1)(b-1))}$$

Guarantee:
family-wise confidence coefficient is at least $(1-\alpha)100\ \%$

Meaning:
in at least $(1-\alpha)100\ \%$ of repetition of experiments, all the intervals in the family cover the true corresponding $L_i's$ $\alpha\ \%$ of repetition of experiments, some tests in the family made false discovery when the null hypothesis was true.

Hypothesis testing (t-test) for individual quantity in this family:

$$H_0^i : L_i = 0 \ H_a^i : L_i \neq 0$$

$$t^* = \frac{\hat{L}_i}{s(\hat{L}_i)}$$

If $|t^*| \leq S$, conclude H_0

If $|t^*| > S$, conclude H_a

Guarantee:
family-wise Type I error is at most α

Meaning:
in at most $\alpha\ \%$ of repetition of experiments, some tests in the family made false discovery when the null hypothesis was true.

Analysis of Factor A (and B) Main Effects (When Factors Do Not Interact)

Multiple Comparison Procedure: Tukey

Suppose we're interested in making inference about all pairwise comparisons of factor level means
i.e., a family containing <u>all pairwise comparisons of factor level means</u>

$$\mathscr{L} = \{D_{ii'} = \mu_i - \mu_{i'} \text{ for } i \neq i'\}$$

$\frac{a(a-1)}{2}$ Pairwise comparisons

$$\hat{D}_{ii'} = \bar{Y}_{i\cdot\cdot} - \bar{Y}_{i'\cdot\cdot} \qquad s^2(\hat{D}_{ii'}) = MSE\frac{2}{b}$$

Analysis of Factor A (and B) Main Effects (When Factors Do Not Interact)

Multiple Comparison Procedure: Tukey

$(1-\alpha)100\,\%$ confidence interval for individual quantity in this family:

$$\hat{D}_{ii'} \pm Ts\left(\hat{D}_{ii'}\right)$$

$$T = \frac{1}{\sqrt{2}} q(1-\alpha; a, (a-1)(b-1))$$

Guarantee:
family-wise confidence coefficient is at least $(1-\alpha)100\,\%$

Meaning:
in at least $(1-\alpha)100\,\%$ of repetition of experiments, all the intervals in the family cover the true corresponding $L_i's$ $\alpha\,\%$ of repetition of experiments, some tests in the family made false discovery when the null hypothesis was true.

Hypothesis testing (t-test) for individual quantity in this family:

$$H_0^i : D_{ii'} = 0 \;\; H_a^i : D_{ii'} = \neq 0$$

$$q^* = \frac{\hat{D}_{ii'}}{s(\hat{D}_{ii'})}$$

If $|q^*| \leq T$, conclude H_0

If $|q^*| > T$, conclude H_a

Guarantee:
family-wise Type I error is at most α

Meaning:
in at most $\alpha\,\%$ of repetition of experiments, some tests in the family made false discovery when the null hypothesis was true.

Example

Make all pairwise comparisons for different sizes of city and regions; use the Bonferroni procedure with a 90 percent family confidence coefficient. State your findings.

There are 3 pairwise comparison for factor A and 1 pairwise comparisons for factor B, 4 in total.

$$B = t(1 - \alpha/8,(a-1)(b-1))$$

```
## [1] 6.205347
```

For factor A, $\hat{D}_{12} = \bar{Y}_{1.} - \bar{Y}_{2.}$

$$\bar{Y}_{1.} - \bar{Y}_{2.} \pm B\sqrt{\frac{2MSE}{b}}$$

Code

```
## [1] -76.97457 313.02543
```

$\hat{D}_{13} = \bar{Y}_{1.} - \bar{Y}_{3.}$

$$\bar{Y}_{1.} - \bar{Y}_{3.} \pm B\sqrt{\frac{2MSE}{a}}$$

Code

```
## [1] -91.97457 328.02543
```

$\hat{D}_{23} = \bar{Y}_{2.} - \bar{Y}_{3.}$

$$\bar{Y}_{2.} - \bar{Y}_{3.} \pm B\sqrt{\frac{2MSE}{b}}$$

Code

```
## [1] -16.97457 403.02543
```

For factor B, $\hat{D}_{12} = \bar{Y}_{.1} - \bar{Y}_{.2}$

$$\bar{Y}_{.1} - \bar{Y}_{.2} \pm B\sqrt{\frac{2MSE}{a}}$$

Code

```
## [1]  28.38777 348.38777
```

For this family of confidence intervals, the following conclusions may be drawn with family confidence coefficient of 90 percent:

- The average premium for different city sizes do not differ

- But the average premium for Eastern cities is higher than Western cities.