Statistical Methods

Nels Grevstad

Metropolitan State University of Denver

ngrevsta@msudenver.edu

October 22, 2019

Topics

1. Two-Factor ANOVA with $K = 1$ (Cont'd)

Objectives

Objectives:

- Carry out Tukey's multiple comparison procedure after a two-factor ANOVA with $K = 1$, and interpret the results.
- Give the definition of a randomized block experiment, state the goal of randomized block experiments and describe their advantage over completely randomized experiments.

Two-Factor ANOVA with $K = 1$ (Cont'd)

Multiple Comparisons in the Additive Effects Model

- After rejecting either H_{0A} or H_{0B}, **Tukey's procedure** can be used to determine *which* **levels** of the factor **differ**.

Tukey's Multiple Comparison Procedure: *After* the two-factor ANOVA F test rejects H_{0A} or H_{0B}:

1. Choose an **overall familywise confidence level** $100(1-\alpha)\%$ (usually $\alpha = 0.05$ for a 95% confidence level).

2. For **Factor A comparisons**, compute the $I(I-1)/2$ **CIs**:

$$\bar{X}_{i\cdot} - \bar{X}_{i'\cdot} \;\pm\; Q_{\alpha,I,IJ-I-J+1}\sqrt{\frac{MSE}{J}}\,.$$

   For **Factor B comparisons**, compute the $J(J-1)/2$ **CIs**:

$$\bar{X}_{\cdot j} - \bar{X}_{\cdot j'} \;\pm\; Q_{\alpha,J,IJ-I-J+1)}\sqrt{\frac{MSE}{I}}\,.$$

3. For any interval that **doesn't contain zero**, deem those **levels** of the given factor to be **different.**

Example

For the study of the effects of **brand** and **storage time** on vitamin C in orange juice, we found that **storage time** had an **effect**, but **brand** didn't.

Example

For the study of the effects of **brand** and **storage time** on vitamin C in orange juice, we found that **storage time** had an **effect**, but **brand** didn't.

The **Tukey procedure** in R produces the following CIs:

Times	Difference	Lower End Pt	Upper End Pt	
Day3-Day7	0.63	-7.779	9.046	
Day0-Day7	10.30	1.887	18.713	*
Day0-Day3	9.67	1.254	18.079	*

Intervals marked with asterisks don't contain zero.

Example

For the study of the effects of **brand** and **storage time** on vitamin C in orange juice, we found that **storage time** had an **effect**, but **brand** didn't.

The **Tukey procedure** in R produces the following CIs:

Times	Difference	Lower End Pt	Upper End Pt	
Day3-Day7	0.63	-7.779	9.046	
Day0-Day7	10.30	1.887	18.713	*
Day0-Day3	9.67	1.254	18.079	*

Intervals marked with asterisks don't contain zero.

We conclude that **Day 0** differs from both **Days 3** and **7**, but **Days 3** and **7** don't differ from each other.

Estimating Parameters in the Additive Effects Model

- Recall that the *additive effects version* of the two-factor ANOVA model is:

$$X_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij} \,. \tag{1}$$

Model Parameter Estimators: We estimate the unknown model parameters μ, α_i, β_j, and σ using the **estimators** $\hat{\mu}$**,** $\hat{\alpha}_i$**,** $\hat{\beta}_j$, and $\hat{\sigma}$ defined as:

Model Parameter	Estimator
μ	$\hat{\mu} = \bar{X}_{..}$
$\alpha_i = \mu_{i.} - \mu$	$\hat{\alpha}_i = \bar{X}_{i.} - \bar{X}_{..}$
$\beta_j = \mu_{.j} - \mu$	$\hat{\beta}_i = \bar{X}_{.j} - \bar{X}_{..}$
σ	$\hat{\sigma} = \sqrt{\text{MSE}}$

Predicted Values and Residuals for the Additive Effects Model

- The ***fitted value*** (or ***predicted value***) for the individual in the i, jth cell, $\hat{X}_{ij}$, is defined as:

$$\begin{aligned} \hat{X}_{ij} &= \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j \\ &= \bar{X}_{..} + (\bar{X}_{i.} - \bar{X}_{..}) + (\bar{X}_{.j} - \bar{X}_{..}) \\ &= \bar{X}_{i.} + \bar{X}_{.j} - \bar{X}_{..} \end{aligned}$$

Predicted Values and Residuals for the Additive Effects Model

- The ***fitted value*** (or ***predicted value***) for the individual in the i, jth cell, $\hat{X}_{ij}$, is defined as:

$$\begin{aligned} \hat{X}_{ij} &= \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j \\ &= \bar{X}_{..} + (\bar{X}_{i.} - \bar{X}_{..}) + (\bar{X}_{.j} - \bar{X}_{..}) \\ &= \bar{X}_{i.} + \bar{X}_{.j} - \bar{X}_{..} \end{aligned}$$

  $\hat{X}_{ij}$ is the value we'd predict, based on the data, for the response of the individual in the i, jth cell.

- The ***residual*** for the observation in the i, jth cell, e_{ij}, is defined as

$$\begin{aligned} e_{ij} &= X_{ij} - \hat{X}_{ij} \\ &= X_{ij} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j) \\ &= X_{ij} - \bar{X}_{i\cdot} - \bar{X}_{\cdot j} + \bar{X}_{\cdot\cdot} \end{aligned}$$

- The ***residual*** for the observation in the i, jth cell, $\boldsymbol{e_{ij}}$, is defined as

$$
\begin{aligned}
e_{ij} &= X_{ij} - \hat{X}_{ij} \\
&= X_{ij} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j) \\
&= X_{ij} - \bar{X}_{i\cdot} - \bar{X}_{\cdot j} + \bar{X}_{\cdot\cdot}
\end{aligned}
$$

The **residual** $\boldsymbol{e_{ij}}$ corresponds to the **random error** term $\boldsymbol{\epsilon_{ij}}$ in the model.

- **Comment**: The **error sum of squares** (Slides 13) is the **sum of squared residuals**, i.e.

$$\text{SSE} = \sum_i \sum_j e_{ij}^2 .$$

Randomized Block Experiments

- In a **one-factor** ***completely randomized experiment***, IJ individuals are randomly split into I treatment groups, with J individuals per group.

Randomized Block Experiments

- In a **one-factor** ***completely randomized experiment***, IJ individuals are randomly split into I treatment groups, with J individuals per group.
- But **heterogeneity** among individuals can inflate the random variation in the observed responses, making it harder to detect treatment effects.

Example

A study investigated the productivity of secretaries with different word processing programs. The study design called for giving an identical task to **nine** secretaries, allocated to **three** treatment groups.

Example

A study investigated the productivity of secretaries with different word processing programs. The study design called for giving an identical task to **nine** secretaries, allocated to **three** treatment groups.

Group 1 used a primarily **menu-driven** program. **Group 2** used a **command-driven** program and **Group 3** used a **mixture** of both approaches.

The time (in minutes) taken to complete the task was recorded.

The secretaries had **different** levels of experience, typing speed, and computer skills.

The secretaries had **different** levels of experience, typing speed, and computer skills.

If a ***completely randomized*** one-factor experiment was carried out, this **heterogeneity** would contribute to the **random variation** in completion times **within** each group.

Factor: Word Processing Program

Menu Driven	Command Driven	Mixture
13	14	11
10	12	8
8	9	7

Factor: Word Processing Program

Menu Driven	Command Driven	Mixture
13	14	11
10	12	8
8	9	7

Some of the observed variation within treatment groups is due to differences in experience levels.

- In a ***randomized block experiment***, the IJ individuals are first divided into J groups of I individuals per group, called ***blocks***, that are **homogeneous** with respect to a so-called ***blocking variable*** that's believed to contribute to variation in observed responses.

- In a ***randomized block experiment***, the IJ individuals are first divided into J groups of I individuals per group, called ***blocks***, that are **homogeneous** with respect to a so-called ***blocking variable*** that's believed to contribute to variation in observed responses.

  Then, separately for each block, the I **individuals within the block** are **randomized** to the I **treatments.**

Example

For the secretary productivity study using a ***randomized block experiment***, the **nine** secretaries are first split into **three *blocks*** (groups) of **three** secretaries each based on **experience level** (less than 1 year, 1 - 5 years, and more than 5 years).

Example

For the secretary productivity study using a ***randomized block experiment***, the **nine** secretaries are first split into **three** ***blocks*** (groups) of **three** secretaries each based on **experience level** (less than 1 year, 1 - 5 years, and more than 5 years).

Then, **within each block**, the **three** secretaries are **randomly assigned** to the three **word processing programs**.

Example

For the secretary productivity study using a ***randomized block experiment***, the **nine** secretaries are first split into **three** ***blocks*** (groups) of **three** secretaries each based on **experience level** (less than 1 year, 1 - 5 years, and more than 5 years).

Then, **within each block**, the **three** secretaries are **randomly assigned** to the three **word processing programs**.

The data are on the next slide.

Factor: Word Processing Program

Blocks: Experience Level		Menu Driven	Command Driven	Mixture
	< 1 Year	13	14	11
	$1 - 5$ Years	10	12	8
	> 5 Years	8	9	7

- In **randomized block experiments**:
  - The effects of the treatments are of major interest to the experimenter.

- In **randomized block experiments**:
  - The effects of the treatments are of major interest to the experimenter.
  - The effects of the blocking variable are generally not of interest.

- In **randomized block experiments**:
  - The effects of the treatments are of major interest to the experimenter.
  - The effects of the blocking variable are generally not of interest.
- The analysis is carried out **exactly** as if the **blocking variable** was a **second factor** in the experiment.

Example

For the study of secretary productivity using the ***randomized blocks design***, the **ANOVA table** is below.

Source of of Variation	**df**	**Sum of Squares**	**Mean Square**	**f**	**P-value**
Blocks (Experience)	2	32.89	16.444	59.2	0.00107
Treatments (Program)	2	13.56	6.778	24.4	0.00574
Error	4	1.11	0.278		
Total	8	47.56			

The **word processing program** has an **statistically significant effect** on the time to complete the task.

- A **randomized block experiment** explicitly **models** the **blocking variable** as a source of **deterministic** (non-random) variation in the data, thereby eliminating it as a contributor to **random variation**.

- A **randomized block experiment** explicitly **models** the **blocking variable** as a source of **deterministic** (non-random) variation in the data, thereby eliminating it as a contributor to **random variation.**

  The goal is to **gain power** for detecting a **treatment effect.**

- A **randomized block experiment** explicitly **models** the **blocking variable** as a source of **deterministic** (non-random) variation in the data, thereby eliminating it as a contributor to **random variation.**

  The goal is to **gain power** for detecting a **treatment effect.**

Example

In the secretary productivity study, **if blocking *wasn't*** used the data would be as shown below.

Factor: Word Processing Program

Menu Driven	Command Driven	Mixture
13	14	11
10	12	8
8	9	7

Example

In the secretary productivity study, **if blocking *wasn't*** used the data would be as shown below.

Factor: Word Processing Program

Menu Driven	Command Driven	Mixture
13	14	11
10	12	8
8	9	7

Some of the **random variation** within groups is due to differences in secretaries' experience levels.

The **one-factor ANOVA table** is below.

Source of of Variation	**df**	**Sum of Squares**	**Mean Square**	**f**	**P-value**
Treatments (Program)	2	13.56	6.778	1.20	0.3650
Error	6	34.00	5.667		
Total	8	47.56			

The **one-factor ANOVA table** is below.

Source of of Variation	**df**	**Sum of Squares**	**Mean Square**	**f**	**P-value**
Treatments (Program)	2	13.56	6.778	1.20	0.3650
Error	6	34.00	5.667		
Total	8	47.56			

The **SSE** here is much **larger** than when **blocking *was*** used.

The **one-factor ANOVA table** is below.

Source of of Variation	df	Sum of Squares	Mean Square	f	P-value
Treatments (Program)	2	13.56	6.778	1.20	0.3650
Error	6	34.00	5.667		
Total	8	47.56			

The **SSE** here is much **larger** than when **blocking *was*** used.

(In fact, the **SSE here** is the the **SSE** for the **blocked model *plus*** the **SSA** for that model.)

The **larger SSE** here leads to a **larger MSE**, **smaller F** value, and **non-significant** treatment (program) **effect.**

- **Comment**: Although blocking leads to a **smaller SSE**, it also leads to **fewer df** for **SSE** ($IJ - I - J + 1$ compared to $I(J-1)$).

- **Comment**: Although blocking leads to a **smaller SSE**, it also leads to **fewer df** for **SSE** ($IJ - I - J + 1$ compared to $I(J-1)$).

  Thus blocking **can** lead to a **larger MSE** if the **reduction** in **SSE** is **small** relative to the decrease in **df**.

- **Comment**: Although blocking leads to a **smaller SSE**, it also leads to **fewer df** for **SSE** ($IJ - I - J + 1$ compared to $I(J - 1)$).

  Thus blocking **can** lead to a **larger MSE** if the **reduction** in **SSE** is **small** relative to the decrease in **df**.

  In this case, there's no advantage to blocking.

- **Comment**: A **matched pairs** study is a ***randomized block experiment*** in which there are ***two*** **treatment groups** and each ***pair*** is a **block.**