10 One-Factor Analysis of Variance (Cont'd)
11 Tests for the Effects of Two Factors

MTH 3240 Environmental Statistics

Spring 2020

Objectives

Objectives:

- Carry out a Bonferroni multiple comparison procedure to identify which of k population means differ from each other (**Optional for Spring 2020**).
- Recognize two-factor studies.

Multiple Comparisons Procedures (Optional for Spring 2020)

- After rejecting the null hypothesis in an ANOVA F or Kruskal-Wallis test, we can determine **which** means differ from each other using a ***multiple comparison*** procedure.

Multiple Comparisons Procedures (Optional for Spring 2020)

- After rejecting the null hypothesis in an ANOVA F or Kruskal-Wallis test, we can determine **which** means differ from each other using a ***multiple comparison*** procedure.
- It can be shown that the total number of comparisons of means is

$$\text{Number of pairs } \mu_i \text{ and } \mu_j \text{ to compare} = \frac{k(k-1)}{2}.$$

(Optional for Spring 2020)

Example

For the lead measurements made at $k = 5$ labs, if we want to know **which** labs differ from each other, we'd need to make

$$\frac{k(k-1)}{2} = \frac{5(5-1)}{2} = 10$$

comparisons, namely

Lab1 vs Lab2
Lab1 vs Lab3
Lab1 vs Lab4
Lab1 vs Lab5
Lab2 vs Lab3
Lab2 vs Lab4
Lab2 vs Lab5
Lab3 vs Lab4
Lab3 vs Lab5

(Optional for Spring 2020)

- We **don't** just perform multiple two-sample t tests (or rank sum tests), each using a level of significance, say, $\alpha = 0.05$.

(Optional for Spring 2020)

- We **don't** just perform multiple two-sample t tests (or rank sum tests), each using a level of significance, say, $\alpha = 0.05$.

  If we did, although the **Type I error** probability on ***any particular*** test would be **0.05**, ...

(Optional for Spring 2020)

- We **don't** just perform multiple two-sample t tests (or rank sum tests), each using a level of significance, say, $\alpha = 0.05$.

  If we did, although the **Type I error** probability on ***any particular*** test would be **0.05**, ...

  the probability of making ***at least one*** **Type I error** among the ***set*** of tests would be substantially **greater** than **0.05**.

(Optional for Spring 2020)

Pairwise and Familywise Type I Error Rates

- Suppose k population means are being tested for differences $\mu_i - \mu_j$ one pair at a time.

(Optional for Spring 2020)

Pairwise and Familywise Type I Error Rates

- Suppose k population means are being tested for differences $\mu_i - \mu_j$ one pair at a time.

  The ***pairwise Type I error rate***, denoted α_p, is the **probability** that any ***particular*** pairwise test will result in a **Type I error.**

(Optional for Spring 2020)

Pairwise and Familywise Type I Error Rates

- Suppose k population means are being tested for differences $\mu_i - \mu_j$ one pair at a time.

  The ***pairwise Type I error rate***, denoted α_p, is the **probability** that any ***particular*** pairwise test will result in a **Type I error.**

  The ***familywise Type I error rate***, denoted α_f, is the **probability** that ***at least one*** of the tests will result in a **Type I error.**

(Optional for Spring 2020)

- The goal in a ***multiple comparison procedure*** is to hold the **familywise Type I error rate** at a fixed level, say **0.05.**

(Optional for Spring 2020)

- The goal in a ***multiple comparison procedure*** is to hold the **familywise Type I error rate** at a fixed level, say **0.05.**
- There are several **multiple comparison procedures.**

(Optional for Spring 2020)

- The goal in a ***multiple comparison procedure*** is to hold the **familywise Type I error rate** at a fixed level, say **0.05**.
- There are several **multiple comparison procedures**.

  We'll look at the simplest one, called the ***Bonferroni procedure***.

(Optional for Spring 2020)

The Bonferroni Procedure

- The **Bonferroni procedure** holds the **familywise Type I error rate** at a fixed level (usually $\alpha_f = 0.05$) by using a sufficiently small level of significance for each pairwise test of hypotheses

$$\begin{aligned} H_0 : \mu_i - \mu_j &= 0 \\ H_a : \mu_i - \mu_j &\neq 0 \end{aligned}$$

(Optional for Spring 2020)

- More specifically, it divides the **familywise Type I** error rate equally among the pairwise tests.

(Optional for Spring 2020)

- More specifically, it divides the **familywise Type I** error rate equally among the pairwise tests.

  Thus, for example, to perform the **10** pairwise tests comparing the **five labs**, we'd use level of significance

  $$\boldsymbol{\alpha_p} = \frac{0.05}{10} = 0.005$$

  for each test.

(Optional for Spring 2020)

Bonferroni Procedure After an ANOVA F Test

- The next slide gives the **Bonferroni procedure** after the null hypothesis is rejected in an **ANOVA F test.**

(Optional for Spring 2020)

Bonferroni Procedure After an ANOVA F Test

- The next slide gives the **Bonferroni procedure** after the null hypothesis is rejected in an **ANOVA** F **test.**

  It merely involves doing **multiple two-sample** t **tests**, but with two adjustments:

(Optional for Spring 2020)

Bonferroni Procedure After an ANOVA F Test

- The next slide gives the **Bonferroni procedure** after the null hypothesis is rejected in an **ANOVA F test.**

  It merely involves doing **multiple two-sample t tests**, but with two adjustments:

  1. We use the **Bonferroni-corrected** level of significance on each test.

(Optional for Spring 2020)

Bonferroni Procedure After an ANOVA F Test

- The next slide gives the **Bonferroni procedure** after the null hypothesis is rejected in an **ANOVA F test.**

  It merely involves doing **multiple two-sample t tests**, but with two adjustments:

  1. We use the **Bonferroni-corrected** level of significance on each test.
  2. We use the **square root** of the **mean squared error** in place of S_i and S_j in the t **test statistics.**

(Optional for Spring 2020)

Bonferroni Multiple Comparison Procedure After One-Factor ANOVA: To decide which pairs of means differ while controlling the familywise Type I error rate at α_f, for each pair of means μ_i and μ_j, test the hypotheses

$$\begin{aligned} H_0 : \mu_i - \mu_j &= 0 \\ H_a : \mu_i - \mu_j &\neq 0 \end{aligned}$$

using the ***Bonferroni pairwise*** t ***test statistic***

$$t = \frac{\bar{Y}_i - \bar{Y}_j - 0}{\sqrt{\frac{\text{MSE}}{n} + \frac{\text{MSE}}{n}}} = \frac{\bar{Y}_i - \bar{Y}_j}{\sqrt{\frac{2 \cdot \text{MSE}}{n}}}$$

(Optional for Spring 2020)

and decision rule

$$\begin{array}{l} \text{Reject } H_0 \text{ if p-value } < \alpha_p \\ \text{Fail to reject } H_0 \text{ if p-value } \geq \alpha_p \text{ .} \end{array}$$

where

$$\alpha_p = \frac{\alpha_f}{(k(k-1)/2)}.$$

When the corresponding H_0 is true, the test statistic t follows a $t(N-k)$ distribution, from which the p-value for that test is obtained.

(Optional for Spring 2020)

Example

For the study of lead measurements at five labs, we'll use the **Bonferroni procedure** to decide ***which*** labs' means differ from each other, while controlling the **familywise Type I error rate** at $\alpha_f = 0.05$.

(Optional for Spring 2020)

Example

For the study of lead measurements at five labs, we'll use the **Bonferroni procedure** to decide ***which*** labs' means differ from each other, while controlling the **familywise Type I error rate** at $\boldsymbol{\alpha_f = 0.05}$.

We need to test **10** sets of hypotheses of the form

$$\begin{aligned} H_0 : \mu_i - \mu_j &= 0 \\ H_a : \mu_i - \mu_j &\neq 0 \end{aligned}$$

(Optional for Spring 2020)

Because $k = 5$, the **Bonferroni-corrected level of significance** to use for each **pairwise test** is

$$\boldsymbol{\alpha_p} = \frac{0.05}{5(5-1)/2} = \mathbf{0.005},$$

(Optional for Spring 2020)

Because $\boldsymbol{k = 5}$, the **Bonferroni-corrected level of significance** to use for each **pairwise test** is

$$\boldsymbol{\alpha_p} = \frac{0.05}{5(5-1)/2} = \mathbf{0.005},$$

and so the decision rule is

Reject H_0 if p-value < 0.005
Fail to reject H_0 if p-value ≥ 0.005

(Optional for Spring 2020)

Statistical software reports the results of **all 10 pairwise tests.** Statistically significant differences (at the Bonferroni-corrected significance level $\alpha_p = 0.005$) are marked with an asterisk.

Pair of Means	t	P-value
Lab1 vs Lab2	1.03	0.3070
Lab1 vs Lab3	-0.50	0.6188
Lab1 vs Lab4	3.69	0.0006*
Lab1 vs Lab5	3.01	0.0043*
Lab2 vs Lab3	-1.53	0.1320
Lab2 vs Lab4	2.66	0.0107
Lab2 vs Lab5	1.97	0.0547
Lab3 vs Lab4	4.20	0.0001*
Lab3 vs Lab5	3.51	0.0010*
Lab4 vs Lab5	-0.69	0.4945

(Optional for Spring 2020)

We conclude that **Labs 1** and **4** differ, **Labs 1** and **5** differ, **Labs 3** and **4** differ, and **Labs 3** and **5** differ.

(Optional for Spring 2020)

Bonferroni Multiple Comparison Procedure After a Kruskal-Wallis Test: To decide which pairs of means differ while controlling the familywise Type I error rate at α_f, for each pair of means μ_i and μ_j, test the hypotheses

$$\begin{aligned} H_0 : \mu_i - \mu_j &= 0 \\ H_a : \mu_i - \mu_j &\neq 0 \end{aligned}$$

using a *rank-sum test* with decision rule

Reject H_0 if p-value $< \alpha_p$
Fail to reject H_0 if p-value $\geq \alpha_p$

where $\alpha_p = \frac{\alpha_f}{(k(k-1)/2)}$.

Two-Factor Studies

- Environmental studies often involve **simultaneously** investigating the effects of **two factors**.

Two-Factor Studies

- Environmental studies often involve **simultaneously** investigating the effects of **two factors**.
- This can involve **samples from populations** or conducting **randomized experiments**.

Two-Factor Studies

- Environmental studies often involve **simultaneously** investigating the effects of **two factors.**
- This can involve **samples from populations** or conducting **randomized experiments.**

  (In the next example, the data are **samples** from **eight populations** defined by **two factors**: **soil type** and **topography.**)

Example

In a study of the effects of **topography** and **soil type** on soil **phosphorus** levels, two **soil types**, *shale-derived* and *sandstone-derived*, were examined in each of four **topographies**: *valleys*, *north-facing slopes*, *south-facing slopes*, and *hilltops*.

Example

In a study of the effects of **topography** and **soil type** on soil **phosphorus** levels, two **soil types**, *shale-derived* and *sandstone-derived*, were examined in each of four **topographies**: *valleys*, *north-facing slopes*, *south-facing slopes*, and *hilltops*.

In each of the **eight** combinations of soil type and topography, **three** phosphorus measurements (ppm) were made, giving a total of **24** phosphorus observations.

The data are shown in the two-way layout below.

Factor B: Topography

		Valley (j=1)	North-Facing (j=2)	South-Facing (j=3)	Hilltop (j=4)	
Factor A: Soil Type	Shale (i=1)	98	78	117	83	
		172	77	54	12	$\bar{Y}_{1\cdot} = 90.5$
		185	100	96	14	
	Sandstone (i=2)	19	27	28	55	
		39	49	53	21	$\bar{Y}_{2\cdot} = 35.9$
		25	24	72	19	
		$\bar{Y}_{\cdot 1} = 89.7$	$\bar{Y}_{\cdot 2} = 59.2$	$\bar{Y}_{\cdot 3} = 70.0$	$\bar{Y}_{\cdot 4} = 34.0$	$\bar{Y} = 63.2$

Here's a summary of the data.

Factor B: Topography

		Valley (j=1)	North-Facing (j=2)	South-Facing (j=3)	Hilltop (j=4)	
Factor A: Soil Type	Shale (i=1)	$\bar{Y}_{11} = 151.7$	$\bar{Y}_{12} = 85.0$	$\bar{Y}_{13} = 89.0$	$\bar{Y}_{14} = 36.3$	$\bar{Y}_{1\cdot} = 90.5$
	Sandstone (i=2)	$\bar{Y}_{21} = 27.7$	$\bar{Y}_{22} = 33.3$	$\bar{Y}_{23} = 51.0$	$\bar{Y}_{24} = 31.7$	$\bar{Y}_{2\cdot} = 35.9$
		$\bar{Y}_{\cdot 1} = 89.7$	$\bar{Y}_{\cdot 2} = 59.2$	$\bar{Y}_{\cdot 3} = 70.0$	$\bar{Y}_{\cdot 4} = 34.0$	$\bar{Y} = 63.2$

The **two *factors*** are **soil type**, which has **two *levels***, and **topography**, which has **four *levels***.

The **two *factors*** are **soil type**, which has **two *levels***, and **topography**, which has **four *levels***.

There were three research questions:

The **two *factors*** are **soil type**, which has **two *levels***, and **topography**, which has **four *levels***.

There were three research questions:

1. Does **soil type** affect phosphorus concentrations?

The **two *factors*** are **soil type**, which has **two *levels***, and **topography**, which has **four *levels***.

There were three research questions:

1. Does **soil type** affect phosphorus concentrations?

2. Does **topography** affect phosphorus concentrations?

The **two *factors*** are **soil type**, which has **two *levels***, and **topography**, which has **four *levels***.

There were three research questions:

1. Does **soil type** affect phosphorus concentrations?
2. Does **topography** affect phosphorus concentrations?
3. If **soil type** has an effect on phosphorus, is the effect ***different*** depending on the **topography**?

These three research questions refer, respectively, to:

These three research questions refer, respectively, to:

1. A **soil type *main effect***.

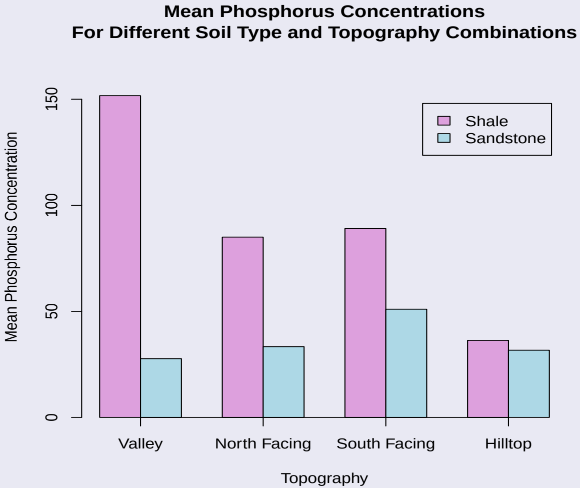
These three research questions refer, respectively, to:

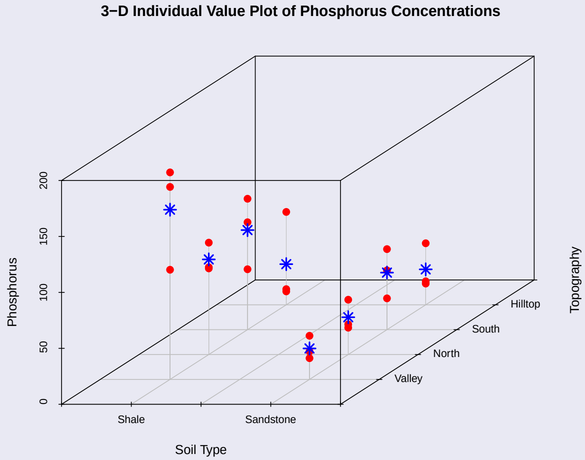
1. A **soil type *main effect***.
2. A **topography *main effect***.

These three research questions refer, respectively, to:

1. A **soil type *main effect***.
2. A **topography *main effect***.
3. An **interaction effect** between **soil type** and **topography**.

Plots of the data are on the next slides.


Mean Phosphorus Concentrations
For Different Soil Type and Topography Combinations
Mean Phosphorus Concentration
0
50
100
150
Shale
Sandstone
Valley
North Facing
South Facing
Hilltop
Topography


3−D Individual Value Plot of Phosphorus Concentrations
Phosphorus
0
50
100
150
200
Shale
Sandstone
Soil Type
Valley
North
South
Hilltop
Topography

Two-Factor ANOVA

- ***Two-factor analysis of variance*** (or **ANOVA**) is a procedure for deciding if either of **two factors** have an effect on a response variable, and if so, whether the effect of one is different depending on the level of the other (i.e. whether there's an **interaction effect**).

- We'll call the factors ***factor*** A and ***factor*** B.

- We'll call the factors ***factor*** A and ***factor*** B.
- In a two-way layout, as in the last example, each **row** corresponds to a ***level*** of **factor** A and each **column** to a ***level*** of **factor** B.

- We'll call the factors ***factor*** A and ***factor*** B.
- In a two-way layout, as in the last example, each **row** corresponds to a ***level*** of **factor** A and each **column** to a ***level*** of **factor** B.
- We'll refer to each of the row-column intersections as a ***group***.

- We'll call the factors ***factor*** A and ***factor*** B.
- In a two-way layout, as in the last example, each **row** corresponds to a ***level*** of **factor** A and each **column** to a ***level*** of **factor** B.
- We'll refer to each of the row-column intersections as a ***group***.

  Each **group** corresponds to a **random sample** of size n from a **population.**

- In practice:
  - Each **group** could also be a **treatment group**, defined by levels of the two factors, in a randomized **experiment.**

- In practice:
  - Each **group** could also be a **treatment group**, defined by levels of the two factors, in a randomized **experiment**.
  - The group sample sizes **don't** all have to be the same, but the notation gets more complicated when they're not.

- **Notation**:

$$\boldsymbol{a} = \text{The number of } \textbf{\textit{levels}} \text{ of Factor } A \text{ (}\textbf{rows}\text{)}$$

$$\boldsymbol{b} = \text{The number of } \textbf{\textit{levels}} \text{ of Factor } B \text{ (}\textbf{columns}\text{)}$$

$$\boldsymbol{n} = \text{The common } \textbf{sample size} \text{ for the } ab \text{ groups.}$$

$$\boldsymbol{Y_{ijk}} = \text{The } k\text{th observation in the } i,j\text{th group.}$$

- **Notation**:

$$\boldsymbol{a} = \text{The number of } \textit{\textbf{levels}} \text{ of Factor } A \text{ } (\textbf{rows})$$

$$\boldsymbol{b} = \text{The number of } \textit{\textbf{levels}} \text{ of Factor } B \text{ } (\textbf{columns})$$

$$\boldsymbol{n} = \text{The common } \textbf{sample size} \text{ for the } ab \text{ groups.}$$

$$\boldsymbol{Y_{ijk}} = \text{The } k\text{th observation in the } i,j\text{th group.}$$

(The first subscript, $\boldsymbol{i}$, indicates the level of Factor A and takes the values $1, 2, \ldots, a$. The second, $\boldsymbol{j}$, indicates the level of Factor B and takes the values $1, 2, \ldots, b$. The third, $\boldsymbol{k}$, distinguishes individuals within a group and takes values $1, 2, \ldots, n$.)

- (cont'd)

$\bar{Y}_{i\cdot}$ = The i***th row mean*** in the two-way layout (or Factor A ***level mean***).

$\bar{Y}_{\cdot j}$ = The j***th column mean*** in the two-way layout (or Factor B ***level mean***).

$\bar{Y}_{ij}$ = The i, j***th group mean***.

N = The ***overall sample size*** for all ab groups combined. Note: $N = abn$.

$\bar{Y}$ = The ***overall sample mean*** of the N observations combined.

Fact: When the group sample sizes are all the same, the overall mean $\bar{Y}$ is equal to all of the following:

1. The average of the ab group means $\bar{Y}_{ij}$.
2. The average of the a row means $\bar{Y}_{1\cdot}, \bar{Y}_{2\cdot}, \ldots, \bar{Y}_{a\cdot}$.
3. The average of the b column means $\bar{Y}_{\cdot 1}, \bar{Y}_{\cdot 2}, \ldots, \bar{Y}_{\cdot b}$.

Example (Cont'd)

For the study of the effects of **topography** and **soil type** on soil **phosphorus**, we have

$$a = 2 \qquad \text{and} \qquad b = 4$$

and also

$$n = 3 \qquad \text{and} \qquad N = 24$$

(The row means, column means, group means, and overall mean are shown in the tables in the last example.)

- A **factor A main effect** is indicated by **variation** in the **row means.**

- A **factor** A **main effect** is indicated by **variation** in the **row means.**

  A **factor** B **main effect** is indicated by **variation** in the **column means.**

- A **factor** A **main effect** is indicated by **variation** in the **row means.**

  A **factor** B **main effect** is indicated by **variation** in the **column means.**

  We refer to these as ***between-rows variation*** and ***between-columns variation***, respectively.

- A **factor** A **main effect** is indicated by **variation** in the **row means.**

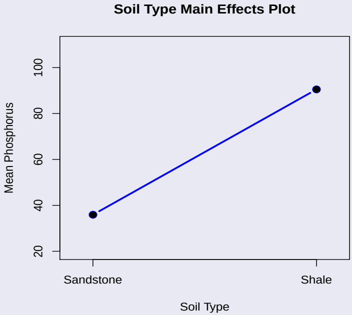
  A **factor** B **main effect** is indicated by **variation** in the **column means.**

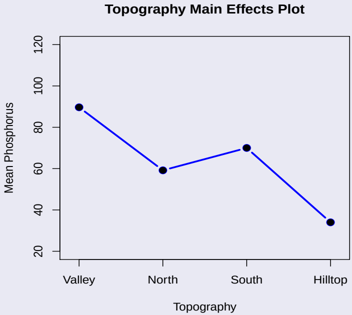
  We refer to these as ***between-rows variation*** and ***between-columns variation***, respectively.

- Variation of individual observations (Y_{ijk}'s) within a group will be referred to as ***within-groups variation***.

We can inspect the **between-rows** and **between-columns variation** in a ***main effects plot*** (or ***level means plot***).

For the soil phosphorus study, the **main effects plots** are below.





- We'll decide if there's a **statistically significant factor A effect** by comparing the **between-rows variation** to **within-groups** variation.

- We'll decide if there's a **statistically significant factor A effect** by comparing the **between-rows variation** to **within-groups** variation.

  We'll decide if there's a **statistically significant factor B effect** by comparing the **between-columns variation** to **within-groups** variation.

- We'll decide if there's a **statistically significant factor A effect** by comparing the **between-rows variation** to **within-groups** variation.

  We'll decide if there's a **statistically significant factor B effect** by comparing the **between-columns variation** to **within-groups** variation.
- (We'll see later how to decide if there's a **statistically significant interaction effect**.)