

Math 362: Mathematical Statistics II

Le Chen

le.chen@emory.edu
chenle02@gmail.com

Emory University
Atlanta, GA

Last updated on Spring 2021
Last compiled on January 15, 2023

2021 Spring

Creative Commons License
(CC By-NC-SA)

Chapter 13. Randomized Block Designs

§ 13.1 Introduction

§ 13.2 The F Test for a Randomized Block Design

§ 13.A Appendix: Some Discussions and Extensions

Plan

§ 13.1 Introduction

§ 13.2 The F Test for a Randomized Block Design

§ 13.A Appendix: Some Discussions and Extensions

Chapter 13. Randomized Block Designs

§ 13.1 Introduction

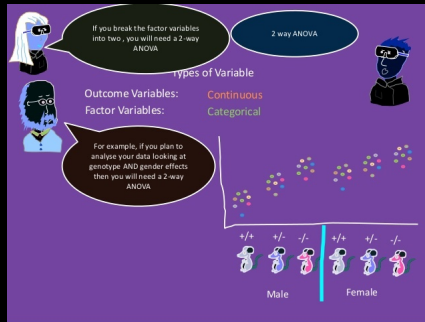
§ 13.2 The F Test for a Randomized Block Design

§ 13.A Appendix: Some Discussions and Extensions

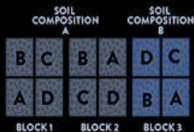
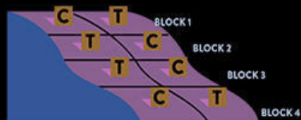
Rationale:

Reducing variability by blocking[†]

[†] *Blocking* is the arranging of experimental units in groups (blocks) that are similar to one another.



<https://www.slideshare.net/KevinHamill2/experimental-design-cartoon-part-5-sample-size>



Goal Reducing variability caused by

a *elevation.*

b *soil types.*

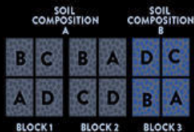
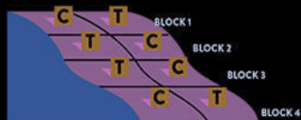
v.s.

c *complete randomized design*

} One-way ANOVA

} Two-way ANOVA

<https://www.sare.org/Learning-Center/Bulletins/How-to-Conduct-Research-on-Your-Farm-or-Ranch/Text-Version/Basics-of-Experimental-Design>



Goal Reducing variability caused by

a *elevation.*

b *soil types.*

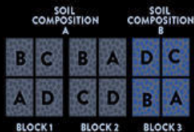
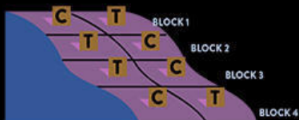
v.s.

c *complete randomized design*

} One-way ANOVA

} Two-way ANOVA

<https://www.sare.org/Learning-Center/Bulletins/How-to-Conduct-Research-on-Your-Farm-or-Ranch/Text-Version/Basics-of-Experimental-Design>



Goal Reducing variability caused by

a *elevation.*

b *soil types.*

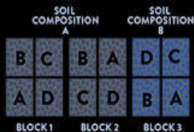
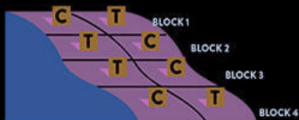
v.s.

c *complete randomized design*

} One-way ANOVA

} Two-way ANOVA

<https://www.sare.org/Learning-Center/Bulletins/How-to-Conduct-Research-on-Your-Farm-or-Ranch/Text-Version/Basics-of-Experimental-Design>



Goal Reducing variability caused by

a *elevation.*

b *soil types.*

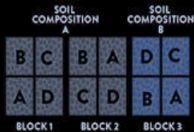
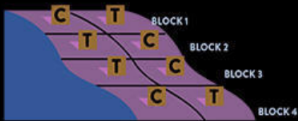
v.s.

c *complete randomized design*

} One-way ANOVA

} Two-way ANOVA

<https://www.sare.org/Learning-Center/Bulletins/How-to-Conduct-Research-on-Your-Farm-or-Ranch/Text-Version/Basics-of-Experimental-Design>



Goal Reducing variability caused by

a *elevation.*

b *soil types.*

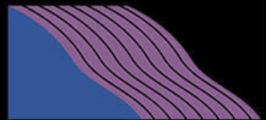
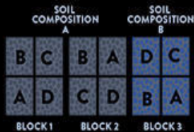
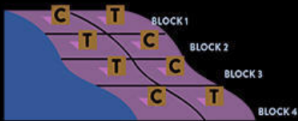
v.s.

c *complete randomized design*

} One-way ANOVA

} Two-way ANOVA

<https://www.sare.org/Learning-Center/Bulletins/How-to-Conduct-Research-on-Your-Farm-or-Ranch/Text-Version/Basics-of-Experimental-Design>



Goal Reducing variability caused by

a *elevation.*

b *soil types.*

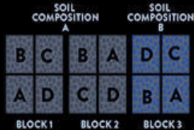
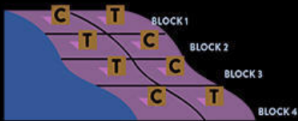
v.s.

c *complete randomized design*

} One-way ANOVA

} Two-way ANOVA

<https://www.sare.org/Learning-Center/Bulletins/How-to-Conduct-Research-on-Your-Farm-or-Ranch/Text-Version/Basics-of-Experimental-Design>



Goal Reducing variability caused by

a *elevation.*

b *soil types.*

v.s.

c *complete randomized design*

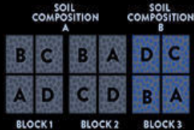
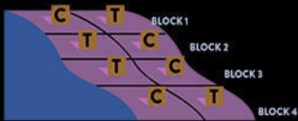
} One-way ANOVA

Two-way ANOVA

<https://www.sare.org/Learning-Center/Bulletins/>

How-to-Conduct-Research-on-Your-Farm-or-Ranch/Text-Version/

Basics-of-Experimental-Design



Goal Reducing variability caused by

a *elevation.*

b *soil types.*

v.s.

c *complete randomized design*

} One-way ANOVA

Two-way ANOVA

<https://www.sare.org/Learning-Center/Bulletins/How-to-Conduct-Research-on-Your-Farm-or-Ranch/Text-Version/Basics-of-Experimental-Design>

Plan

§ 13.1 Introduction

§ 13.2 The F Test for a Randomized Block Design

§ 13.A Appendix: Some Discussions and Extensions

Chapter 13. Randomized Block Designs

§ 13.1 Introduction

§ 13.2 The F Test for a Randomized Block Design

§ 13.A Appendix: Some Discussions and Extensions

Setup Y_{ij} indep. $\sim N(\mu_j + \beta_i, \sigma^2)$, i.e., $Y_{ij} = \mu_j + \beta_i + \epsilon_{ij}$, ϵ_{ij} i.i.d. $\sim N(0, \sigma^2)$

Setup Y_{ij} indep. $\sim N(\mu_j + \beta_i, \sigma^2)$, i.e., $Y_{ij} = \mu_j + \beta_i + \epsilon_{ij}$, ϵ_{ij} i.i.d. $\sim N(0, \sigma^2)$

Setup Y_{ij} indep. $\sim N(\mu_j + \beta_i, \sigma^2)$, i.e., $Y_{ij} = \mu_j + \beta_i + \epsilon_{ij}$, ϵ_{ij} i.i.d. $\sim N(0, \sigma^2)$

Table 13.2.1								
	Treatment Level					<i>Block Totals</i>	<i>Block Means</i>	<i>True Block Effects</i>
		<i>1</i>	<i>2</i>	...	<i>k</i>			
<i>Blocks</i>	1	Y_{11}	Y_{12}	...	Y_{1k}	$T_{1.}$	$\bar{Y}_{1.}$	β_1
	2	Y_{21}	Y_{22}		Y_{2k}	$T_{2.}$	$\bar{Y}_{2.}$	β_2
	\vdots	\vdots			\vdots	\vdots	\vdots	\vdots
	<i>b</i>	Y_{b1}	Y_{b2}		Y_{bk}	$T_{b.}$	$\bar{Y}_{b.}$	β_b
Sample totals		$T_{.1}$	$T_{.2}$		$T_{.k}$	$T_{..}$		
Sample means		$\bar{Y}_{.1}$	$\bar{Y}_{.2}$...	$\bar{Y}_{.k}$		$\bar{Y}_{..}$	
True means		μ_1	μ_2		μ_k			

Recall For one-way ANOVA,

$$Y_{ij} = \mu_j + \epsilon_{ij}$$

\Downarrow

$$\begin{aligned} SSTOT &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left[\left(Y_{ij} - \bar{Y}_{.j} \right) + \left(\bar{Y}_{.j} - \bar{Y}_{..} \right) \right]^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{.j} \right)^2 + \text{zero cross term} + \sum_{i=1}^b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{.j} \right)^2 + b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 \\ &= SSE + SSTR \end{aligned}$$

Recall For one-way ANOVA,

$$Y_{ij} = \mu_j + \epsilon_{ij}$$

\Downarrow

$$\begin{aligned} SSTOT &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left[\left(Y_{ij} - \bar{Y}_{.j} \right) + \left(\bar{Y}_{.j} - \bar{Y}_{..} \right) \right]^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{.j} \right)^2 + \text{zero cross term} + \sum_{i=1}^b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{.j} \right)^2 + b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 \\ &= SSE + SSTR \end{aligned}$$

Recall For one-way ANOVA,

$$Y_{ij} = \mu_j + \epsilon_{ij}$$

\Downarrow

$$\begin{aligned} SSTOT &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left[\left(Y_{ij} - \bar{Y}_{.j} \right) + \left(\bar{Y}_{.j} - \bar{Y}_{..} \right) \right]^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{.j} \right)^2 + \text{zero cross term} + \sum_{i=1}^b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{.j} \right)^2 + b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 \\ &= SSE + SSTR \end{aligned}$$

$$SSTOT = SSE + SSTR$$

$$\Downarrow$$

$$\frac{SSTOT}{\sigma^2} = \frac{SSE}{\sigma^2} + \frac{SSTR}{\sigma^2}$$

$$\wr$$

$$\wr$$

$$\wr$$

$$\chi^2(bk - 1) \quad \chi^2(bk - k) \quad \perp \quad \chi^2(k - 1)$$

Under H_0

✓

Under H_0

$$H_0 : \mu_1 = \cdots = \mu_k$$

Symmetry If

$$Y_{ij} = \beta_i + \epsilon_{ij}$$

\Downarrow

$$\begin{aligned} SSTOT &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left[\left(Y_{ij} - \bar{Y}_{i.} \right) + \left(\bar{Y}_{i.} - \bar{Y}_{..} \right) \right]^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{i.} \right)^2 + \text{zero cross term} + \sum_{i=1}^b \sum_{j=1}^k \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{i.} \right)^2 + k \sum_{i=1}^b \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 \\ &= SSE + SSB \end{aligned}$$

Symmetry If

$$Y_{ij} = \beta_i + \epsilon_{ij}$$

\Downarrow

$$\begin{aligned} SSTOT &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left[\left(Y_{ij} - \bar{Y}_{i.} \right) + \left(\bar{Y}_{i.} - \bar{Y}_{..} \right) \right]^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{i.} \right)^2 + \text{zero cross term} + \sum_{i=1}^b \sum_{j=1}^k \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{i.} \right)^2 + k \sum_{i=1}^b \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 \\ &= SSE + SSB \end{aligned}$$

Symmetry If

$$Y_{ij} = \beta_i + \epsilon_{ij}$$

\Downarrow

$$\begin{aligned} SSTOT &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left[\left(Y_{ij} - \bar{Y}_{i.} \right) + \left(\bar{Y}_{i.} - \bar{Y}_{..} \right) \right]^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{i.} \right)^2 + \text{zero cross term} + \sum_{i=1}^b \sum_{j=1}^k \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{i.} \right)^2 + k \sum_{i=1}^b \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 \\ &= SSE + SSB \end{aligned}$$

$$SSTOT = SSE + SSB$$

$$\Downarrow$$

$$\frac{SSTOT}{\sigma^2} = \frac{SSE}{\sigma^2} + \frac{SSB}{\sigma^2}$$

$$\downarrow$$

$$\downarrow$$

$$\downarrow$$

$$\chi^2(bk - 1) \quad \chi^2(bk - b) \quad \perp \quad \chi^2(b - 1)$$

Under \tilde{H}_0

✓

Under \tilde{H}_0

$$\tilde{H}_0 : \beta_1 = \cdots = \beta_b$$

Similarly If

$$Y_{ij} = \mu_j + \beta_i + \epsilon_{ij}$$

\Downarrow

$$\begin{aligned} SSTOT &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left[\left(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} \right) + \left(\bar{Y}_{i.} - \bar{Y}_{..} \right) + \left(\bar{Y}_{.j} - \bar{Y}_{..} \right) \right]^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} \right)^2 + \sum_{i=1}^b \sum_{j=1}^k \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 \\ &\quad + \sum_{i=1}^b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 + \text{zero cross terms} \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} \right)^2 + k \sum_{i=1}^b \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 + b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 \\ &= SSE + SSB + SSTR \end{aligned}$$

Similarly If

$$Y_{ij} = \mu_j + \beta_i + \epsilon_{ij}$$

\Downarrow

$$\begin{aligned} SSTOT &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left[\left(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} \right) + \left(\bar{Y}_{i.} - \bar{Y}_{..} \right) + \left(\bar{Y}_{.j} - \bar{Y}_{..} \right) \right]^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} \right)^2 + \sum_{i=1}^b \sum_{j=1}^k \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 \\ &\quad + \sum_{i=1}^b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 + \text{zero cross terms} \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} \right)^2 + k \sum_{i=1}^b \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 + b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 \\ &= SSE + SSB + SSTR \end{aligned}$$

Similarly If

$$Y_{ij} = \mu_j + \beta_i + \epsilon_{ij}$$

\Downarrow

$$\begin{aligned} SSTOT &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{..} \right)^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left[\left(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} \right) + \left(\bar{Y}_{i.} - \bar{Y}_{..} \right) + \left(\bar{Y}_{.j} - \bar{Y}_{..} \right) \right]^2 \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} \right)^2 + \sum_{i=1}^b \sum_{j=1}^k \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 \\ &\quad + \sum_{i=1}^b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 + \text{zero cross terms} \\ &= \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} \right)^2 + k \sum_{i=1}^b \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 + b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 \\ &= SSE + SSB + SSTR \end{aligned}$$

$$SSTOT = SSE + SSB + SSTR$$

↓

$$\frac{SSTOT}{\sigma^2} = \frac{SSE}{\sigma^2} + \frac{SSB}{\sigma^2} + \frac{SSTR}{\sigma^2}$$

}

}

}

}

$$\chi^2(bk - 1) \quad \chi^2((k - 1)(b - 1)) \perp \chi^2(b - 1) \perp \chi^2(k - 1)$$

Under H_0 or \tilde{H}_0

✓

under \tilde{H}_0

under H_0

$$\tilde{H}_0 : \beta_1 = \cdots = \beta_b \quad \text{and} \quad H_0 : \mu_1 \cdots = \mu_k$$

$$H_0 : \mu_1 \cdots = \mu_k$$

↓

Table 13.2.2					
Source	df	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>P</i>
Treatments	$k - 1$	<i>SSTR</i>	$SSTR/(k - 1)$	$\frac{SSTR/(k - 1)}{SSE/(b - 1)(k - 1)}$	$P[F_{k-1, (b-1)(k-1)} \geq \text{obs. } F]$
Blocks	$b - 1$	<i>SSB</i>	$SSB/(b - 1)$	$\frac{SSB/(b - 1)}{SSE/(b - 1)(k - 1)}$	$P[F_{b-1, (b-1)(k-1)} \geq \text{obs. } F]$
Error	$(b - 1)(k - 1)$	<i>SSE</i>	$SSE/(b - 1)(k - 1)$		
Total	$n - 1$	<i>SSTOT</i>			

↑

$$\tilde{H}_0 : \beta_1 = \cdots = \beta_b$$

Computing formulas

$$C = \frac{T_{..}^2}{bk}$$

$$\textcolor{red}{SSTR} = b \sum_{j=1}^k \left(\bar{Y}_{.j} - \bar{Y}_{..} \right)^2 = b \sum_{j=1}^k \bar{Y}_{.j}^2 - bk \bar{Y}_{..}^2 = \frac{1}{b} \sum_{j=1}^k \bar{T}_{.j}^2 - C.$$

$$\textcolor{teal}{SSB} = k \sum_{i=1}^b \left(\bar{Y}_{i.} - \bar{Y}_{..} \right)^2 = k \sum_{i=1}^b \bar{Y}_{i.}^2 - bk \bar{Y}_{..}^2 = \frac{1}{k} \sum_{i=1}^b \bar{T}_{i.}^2 - C.$$

$$SSTOT = \sum_{i=1}^b \sum_{j=1}^k \left(Y_{ij} - \bar{Y}_{..} \right)^2 = \sum_{i=1}^b \sum_{j=1}^k Y_{ij}^2 - bk \bar{Y}_{..}^2 = \sum_{i=1}^b \sum_{j=1}^k Y_{ij}^2 - C.$$

$$SSE = SSTOT - \textcolor{teal}{SSB} - \textcolor{red}{SSTR}$$

E.g. Two methods to test wines: whether these two procedures produce the same measurements?

	DRS-FTIR	Standard
White wine 1	112.9	115.1
White wine 2	123.1	125.6
Red wine 1	135.2	132.4
Red wine 2	140.2	143.7

Test at $\alpha = 0.05$

$$H_0 : \mu_{DRS} = \mu_{STD} \quad \text{v.s.} \quad H_1 : \mu_{DRS} \neq \mu_{STD}$$

and

$$\tilde{H}_0 : \mu_{W1} = \mu_{W2} = \mu_{R1} = \mu_{R2} \quad \text{v.s.} \quad \tilde{H}_1 : \text{not equal}$$

E.g. Two methods to test wines: whether these two procedures produce the same measurements?

	DRS-FTIR	Standard
White wine 1	112.9	115.1
White wine 2	123.1	125.6
Red wine 1	135.2	132.4
Red wine 2	140.2	143.7

Test at $\alpha = 0.05$

$$H_0 : \mu_{DRS} = \mu_{STD} \quad \text{v.s.} \quad H_1 : \mu_{DRS} \neq \mu_{STD}$$

and

$$\tilde{H}_0 : \mu_{W1} = \mu_{W2} = \mu_{R1} = \mu_{R2} \quad \text{v.s.} \quad \tilde{H}_1 : \text{not equal}$$

```

1 > # Case Study 13.2.1
2 > # install .packages("ggpubr")
3 > DIRS <- c(112.9, 123.1, 135.2, 140.2)
4 > STD <- c(115.1, 125.6, 132.4, 143.7)
5 > Wines <- c("W1", "W2", "R1", "R2")
6 > # Create a data frame
7 > my_data <- data.frame(
8 +   method = rep(c("DIRS", "STD"), each =4),
9 +   types = c(Wines,Wines),
10 +   concentration = c(DIRS, STD)
11 + )
12 > # Show data
13 > print(my_data)
14   method types concentration
15 1   DIRS   W1         112.9
16 2   DIRS   W2         123.1
17 3   DIRS   R1         135.2
18 4   DIRS   R2         140.2
19 5     STD   W1         115.1
20 6     STD   W2         125.6
21 7     STD   R1         132.4
22 8     STD   R2         143.7

```

```

1 > # Compute t-test with equal variances
2 > res <- t.test(concentration ~ method,
3 +               data = my_data,
4 +               var.equal = TRUE)
5 > res
6
7 Two Sample t-test
8
9 data: concentration by method
10 t = -0.15721, df = 6, p-value = 0.8802
11 alternative hypothesis: true difference in
    means is not equal to 0
12 95 percent confidence interval:
13  -22.362 19.662
14 sample estimates:
15 mean in group DIRS mean in group STD
16      127.85      129.20

```

```

1 > # The following one-way ANOVA is
    equivalent
2 > # to the two-sample t test
3 > library(car)
4 > model3 = lm(concentration ~ method,
5 +             data=my_data)
6 > Anova(model3)
7 Anova Table (Type II tests)
8
9 Response: concentration
10      Sum Sq Df F value Pr(>F)
11 method    3.64  1  0.0247 0.8802
12 Residuals 884.87 6

```

```

1 > # Compute t-test with unequal variances
2 > res <- t.test(concentration ~ method,
3 +               data = my_data,
4 +               var.equal = FALSE)
5 > res
6
7 Welch Two Sample t-test
8
9 data: concentration by method
10 t = -0.15721, df = 5.9968, p-value = 0.8802
11 alternative hypothesis: true difference in
    means is not equal to 0
12 95 percent confidence interval:
13  -22.3647 19.6647
14 sample estimates:
15 mean in group DIRS mean in group STD
16      127.85      129.20

```

1. Classical method
2. Welch approximation
3. one-way ANOVA



The same answer
(p-value)

Concl. Fail to reject H_0

```

1 > # Compute t-test with equal variances
2 > res <- t.test(concentration ~ method,
3 +               data = my_data,
4 +               var.equal = TRUE)
5 > res
6
7 Two Sample t-test
8
9 data: concentration by method
10 t = -0.15721, df = 6, p-value = 0.8802
11 alternative hypothesis: true difference in
12 means is not equal to 0
13 95 percent confidence interval:
14 -22.362 19.662
15 sample estimates:
16 mean in group DIRS mean in group STD
17 127.85 129.20

```

```

1 > # The following one-way ANOVA is
2   # to the two-sample t test
3 > library(car)
4 > model3 = lm(concentration ~ method,
5 +             data=my_data)
6 > Anova(model3)
7 Anova Table (Type II tests)
8
9 Response: concentration
10 Sum Sq Df F value Pr(>F)
11 method 3.64 1 0.0247 0.8802
12 Residuals 884.87 6

```

```

1 > # Compute t-test with unequal variances
2 > res <- t.test(concentration ~ method,
3 +               data = my_data,
4 +               var.equal = FALSE)
5 > res
6
7 Welch Two Sample t-test
8
9 data: concentration by method
10 t = -0.15721, df = 5.9968, p-value = 0.8802
11 alternative hypothesis: true difference in
12 means is not equal to 0
13 95 percent confidence interval:
14 -22.3647 19.6647
15 sample estimates:
16 mean in group DIRS mean in group STD
17 127.85 129.20

```

1. Classical method
2. Welch approximation
3. one-way ANOVA



The same answer
(p-value)

Concl. Fail to reject H_0

```

1 > # Compute t-test with equal variances
2 > res <- t.test(concentration ~ method,
3 +               data = my_data,
4 +               var.equal = TRUE)
5 > res
6
7 Two Sample t-test
8
9 data: concentration by method
10 t = -0.15721, df = 6, p-value = 0.8802
11 alternative hypothesis: true difference in
12 means is not equal to 0
13 95 percent confidence interval:
14 -22.362 19.662
15 sample estimates:
16 mean in group DIRS mean in group STD
17 127.85 129.20

```

```

1 > # The following one-way ANOVA is
2 # equivalent
3 > # to the two-sample t test
4 > library(car)
5 > model3 = lm(concentration ~ method,
6 +             data=my_data)
7 > Anova(model3)
8 Anova Table (Type II tests)
9
10 Response: concentration
11 Sum Sq Df F value Pr(>F)
12 method 3.64 1 0.0247 0.8802
13 Residuals 884.87 6

```

```

1 > # Compute t-test with unequal variances
2 > res <- t.test(concentration ~ method,
3 +               data = my_data,
4 +               var.equal = FALSE)
5 > res
6
7 Welch Two Sample t-test
8
9 data: concentration by method
10 t = -0.15721, df = 5.9968, p-value = 0.8802
11 alternative hypothesis: true difference in
12 means is not equal to 0
13 95 percent confidence interval:
14 -22.3647 19.6647
15 sample estimates:
16 mean in group DIRS mean in group STD
17 127.85 129.20

```

1. Classical method
2. Welch approximation
3. one-way ANOVA



The same answer
(p-value)

Concl. Fail to reject H_0

```

1 > # Compute t-test with equal variances
2 > res <- t.test(concentration ~ method,
3 +               data = my_data,
4 +               var.equal = TRUE)
5 > res
6
7 Two Sample t-test
8
9 data: concentration by method
10 t = -0.15721, df = 6, p-value = 0.8802
11 alternative hypothesis: true difference in
12 means is not equal to 0
13 95 percent confidence interval:
14 -22.362 19.662
15 sample estimates:
16 mean in group DIRS mean in group STD
17 127.85 129.20

```

```

1 > # The following one-way ANOVA is
2   # to the two-sample t test
3 > library(car)
4 > model3 = lm(concentration ~ method,
5 +             data=my_data)
6 > Anova(model3)
7 Anova Table (Type II tests)
8
9 Response: concentration
10 Sum Sq Df F value Pr(>F)
11 method 3.64 1 0.0247 0.8802
12 Residuals 884.87 6

```

```

1 > # Compute t-test with unequal variances
2 > res <- t.test(concentration ~ method,
3 +               data = my_data,
4 +               var.equal = FALSE)
5 > res
6
7 Welch Two Sample t-test
8
9 data: concentration by method
10 t = -0.15721, df = 5.9968, p-value = 0.8802
11 alternative hypothesis: true difference in
12 means is not equal to 0
13 95 percent confidence interval:
14 -22.3647 19.6647
15 sample estimates:
16 mean in group DIRS mean in group STD
17 127.85 129.20

```

1. Classical method
2. Welch approximation
3. one-way ANOVA



The same answer
(p-value)

Concl. Fail to reject H_0

```

1 > # Now let's carry out two-way ANOVA
2 > library(car)
3 > model = lm(concentration ~ method + types,
4 +           data=my_data)
5 > Anova(model)
6 Anova Table (Type II tests)
7
8 Response: concentration
9             Sum Sq Df F value  Pr(>F)
10 method      3.65   1  0.9154 0.409258
11 types      872.92   3 73.0787 0.002652 **
12 Residuals  11.94   3

```

```

1 > # Now let's try one-way ANOVA
2 > model2 = lm(concentration ~ types,
3 +           data=my_data)
4 > Anova(model2)
5 Anova Table (Type II tests)
6
7 Response: concentration
8             Sum Sq Df F value  Pr(>F)
9 types      872.92   3  74.657 0.0005739 ***
10 Residuals  15.59   4
11 ---
12 Signif. codes:  0 '***' 0.001 '**' 0.01 '*'
                  0.05 '.' 0.1 ' ' 1

```

1. Fail to reject H_0

2. Reject \widetilde{H}_0


```

1 > # Now let's carry out two-way ANOVA
2 > library(car)
3 > model = lm(concentration ~ method + types,
4 +           data=my_data)
5 > Anova(model)
6 Anova Table (Type II tests)
7
8 Response: concentration
9           Sum Sq Df F value  Pr(>F)
10 method      3.65  1  0.9154 0.409258
11 types      872.92  3 73.0787 0.002652 **
12 Residuals  11.94  3

```

```

1 > # Now let's try one-way ANOVA
2 > model2 = lm(concentration ~ types,
3 +           data=my_data)
4 > Anova(model2)
5 Anova Table (Type II tests)
6
7 Response: concentration
8           Sum Sq Df F value  Pr(>F)
9 types      872.92  3  74.657 0.0005739 ***
10 Residuals  15.59  4
11 ---
12 Signif. codes:  0 '***' 0.001 '**' 0.01 '*'
                  0.05 '.' 0.1 ' ' 1

```

1. Fail to reject H_0
2. Reject \tilde{H}_0

E.g. 2 https://rcompanion.org/rcompanion/d_08.html

Test at $\alpha = 0.05$

$$H_0 : \mu_F = \mu_M \quad v.s. \quad H_1 : \mu_F \neq \mu_M$$

and

$$\tilde{H}_0 : \mu_{FF} = \mu_S = \mu_{SS} \quad v.s. \quad \tilde{H}_1 : \text{not all equal}$$

E.g. 2 https://rcompanion.org/rcompanion/d_08.html

Genotype	Female	Male
FF	2.838	1.884
	4.216	2.283
	2.889	4.939
	4.198	3.486
FS	3.550	2.396
	4.556	2.956
	3.087	3.105
	1.943	2.649
SS	3.620	2.801
	3.079	3.421
	3.586	4.275
	2.669	3.110

Test at $\alpha = 0.05$

$$H_0 : \mu_F = \mu_M \quad v.s. \quad H_1 : \mu_F \neq \mu_M$$

and

$$\tilde{H}_0 : \mu_{FF} = \mu_S = \mu_{SS} \quad v.s. \quad \tilde{H}_1 : \text{not all equal}$$

1	> Data				
2		id	Sex	Genotype	Activity
3	1	1	male	ff	1.884
4	2	2	male	ff	2.283
5	3	3	male	fs	2.396
6	4	4	female	ff	2.838
7	5	5	male	fs	2.956
8	6	6	female	ff	4.216
9	7	7	female	ss	3.620
10	8	8	female	ff	2.889
11	9	9	female	fs	3.550
12	10	10	male	fs	3.105
13	11	11	female	fs	4.556
14	12	12	female	fs	3.087
15	13	13	male	ff	4.939
16	14	14	male	ff	3.486
17	15	15	female	ss	3.079
18	16	16	male	fs	2.649

1	17	17	female	fs	1.943
2	18	19	female	ff	4.198
3	19	20	female	ff	2.473
4	20	22	female	ff	2.033
5	21	24	female	fs	2.200
6	22	25	female	fs	2.157
7	23	26	male	ss	2.801
8	24	28	male	ss	3.421
9	25	29	female	ff	1.811
10	26	30	female	fs	4.281
11	27	32	female	fs	4.772
12	28	34	female	ss	3.586
13	29	36	female	ff	3.944
14	30	38	female	ss	2.669
15	31	39	female	ss	3.050
16	32	41	male	ss	4.275
17	33	43	female	ss	2.963
18	34	46	female	ss	3.236
19	35	48	female	ss	3.673
20	36	49	male	ss	3.110

```

1 > # Two-way ANOVA
2 > model = lm(Activity ~ Sex + Genotype,
3 +           data=Data)
4 > Anova(model, type="II")
5 Anova Table (Type II tests)
6
7 Response: Activity
8           Sum Sq Df F value Pr(>F)
9 Sex         0.0681  1  0.0888 0.7676
10 Genotype   0.2772  2  0.1808 0.8354
11 Residuals 24.5285 32
12 > # One-way ANOVA
13 > model_Sex = lm(Activity ~ Sex,
14 +               data=Data)
15 > Anova(model_Sex, type="II")
16 Anova Table (Type II tests)
17
18 Response: Activity
19           Sum Sq Df F value Pr(>F)
20 Sex         0.0681  1  0.0933 0.7619
21 Residuals 24.8057 34
22 > # One-way ANOVA
23 > model_Genotype = lm(Activity ~ Genotype,
24 +                     data=Data)
25 > Anova(model_Genotype, type="II")
26 Anova Table (Type II tests)
27
28 Response: Activity
29           Sum Sq Df F value Pr(>F)
30 Genotype   0.2772  2  0.186 0.8312
31 Residuals 24.5965 33

```

Tukey's pairwise comparison

Replace $Q_{\alpha,k,b(k-k)}$ by $Q_{\alpha,k,(b-1)(k-1)}$

```

1 > # Tukey's pairwise comparison (One-way)
2 > model1 = aov(Activity ~ Genotype,
3 +             data=Data)
4 > TukeyHSD(model1, "Genotype", ordered =
5   TRUE)
6   Tukey multiple comparisons of means
7   95% family-wise confidence level
8   factor levels have been ordered
9   Fit : aov(formula = Activity ~ Genotype, data
10  = Data)
11 $Genotype
12      diff      lwr      upr      p
      adj
13 fs-ff 0.05483333 -0.8100204 0.919687
14      0.9867505
15 ss-ff 0.20741667 -0.6574370 1.072270
16      0.8272105
17 ss-fs 0.15258333 -0.7122704 1.017437
18      0.9021607
  
```

```

1 > # Tukey's pairwise comparison (Two-way)
2 > model2 = aov(Activity ~ Sex + Genotype,
3 +             data=Data)
4 > TukeyHSD(model2, "Genotype", ordered =
5   TRUE)
6   Tukey multiple comparisons of means
7   95% family-wise confidence level
8   factor levels have been ordered
9   Fit : aov(formula = Activity ~ Sex +
10  Genotype, data = Data)
11 $Genotype
12      diff      lwr      upr      p
      adj
13 fs-ff 0.05483333 -0.8234920 0.9331586
14      0.987114
15 ss-ff 0.20741667 -0.6709086 1.0857420
16      0.831554
17 ss-fs 0.15258333 -0.7257420 1.0309086
18      0.904729
  
```

Remark By two-way ANOVA, or through blocking one factor, we obtain

1. larger p -values:
more conservative to reject H_0 .
2. wider C.I.'s:
more conservative on our estimates.

Remark By two-way ANOVA, or through blocking one factor, we obtain

1. larger p -values:

more conservative to reject H_0 .

2. wider C.I.'s:

more conservative on our estimates.

Remark By two-way ANOVA, or through blocking one factor, we obtain

1. larger p -values:
more conservative to reject H_0 .
2. wider C.I.'s:
more conservative on our estimates.

Remark By two-way ANOVA, or through blocking one factor, we obtain

1. larger p -values:
more conservative to reject H_0 .
2. wider C.I.'s:
more conservative on our estimates.

Remark By two-way ANOVA, or through blocking one factor, we obtain

1. larger p -values:
more conservative to reject H_0 .
2. wider C.I.'s:
more conservative on our estimates.

Plan

§ 13.1 Introduction

§ 13.2 The F Test for a Randomized Block Design

§ 13.A Appendix: Some Discussions and Extensions

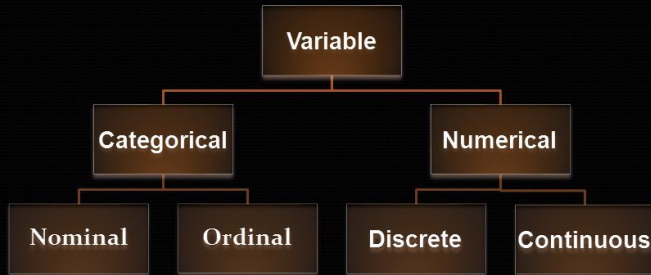
Chapter 13. Randomized Block Designs

§ 13.1 Introduction

§ 13.2 The F Test for a Randomized Block Design

§ 13.A Appendix: Some Discussions and Extensions

Types of Variable with respect to data



Gender

1. Male
2. Female

Motivation

1. Highly Motivated
2. Moderately Motivated
3. Less Motivated

1. No of students
2. No of chairs
3. Collar size

1. Height
2. Weight
3. speed

	Numerical Values	Categorical Values
n Samples	Sample 1 $y_{11} \dots y_{1m}$	$c_{11} \dots c_{1p}$
	Sample 2 $y_{21} \dots y_{2m}$	$c_{21} \dots c_{2p}$
	\vdots	\vdots
	Sample n $y_{n1} \dots y_{nm}$	$c_{n1} \dots c_{np}$
	m	p

		Dependent variable	
Independent variable		Continuous	Categorical
	Categorical	Linear Regression	Logistic Regression
	Continuous	t-test/ANOVA	Chi-square test

Indep. v.s. Dependent

1. Categorical v.s. Continuous

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

MANOVA refers to the multivariate analysis of variance.

ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

ANOVA: Analysis of Variance

Indep. v.s. Dependent

1. Categorical v.s. Continuous

ANOVA

MANOVA

ANCOVA

ANCOVA

ANCOVA

ANCOVA

ANCOVA

ANCOVA

MANOVA refers to the multivariate analysis of variance.

ANCOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

ANCOVA

ANCOVA

ANCOVA

Indep. v.s. Dependent

1. Categorical v.s. Continuous

1.1 $p = 1, m = 1,$	One-way ANOVA
1.2 $p = 2, m = 1,$	Two-way ANOVA
1.3 $p \geq 3, m = 1,$	p -way ANOVA
1.4 $p = 1, m \geq 2,$	One-way MANOVA ^a
1.5 $p = 2, m \geq 2,$	Two-way MANOVA
1.6 $p \geq 3, m \geq 2,$	p -way ANOVA

^aMANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

Correlation
Regression
ANOVA
MANOVA

Indep. v.s. Dependent

1. Categorical v.s. Continuous

1.1 $p = 1, m = 1,$

One-way ANOVA

1.2 $p = 2, m = 1,$

Two-way ANOVA

1.3 $p \geq 3, m = 1,$

p -way ANOVA

1.4 $p = 1, m \geq 2,$

One-way MANOVA^a

1.5 $p = 2, m \geq 2,$

Two-way MANOVA

1.6 $p \geq 3, m \geq 2,$

p -way ANOVA

^aMANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

Two-way ANOVA
Two-way MANOVA
Two-way ANCOVA
Two-way MANCOVA

Indep. v.s. Dependent

1. Categorical v.s. Continuous

1.1 $p = 1, m = 1,$

One-way ANOVA

1.2 $p = 2, m = 1,$

Two-way ANOVA

1.3 $p \geq 3, m = 1,$

p -way ANOVA

1.4 $p = 1, m \geq 2,$

One-way MANOVA^a

1.5 $p = 2, m \geq 2,$

Two-way MANOVA

1.6 $p \geq 3, m \geq 2,$

p -way ANOVA

^aMANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

Correlation
Regression
ANOVA
MANOVA

Indep. v.s. Dependent

1. Categorical v.s. Continuous

1.1 $p = 1, m = 1,$

One-way ANOVA

1.2 $p = 2, m = 1,$

Two-way ANOVA

1.3 $p \geq 3, m = 1,$

p -way ANOVA

1.4 $p = 1, m \geq 2,$

One-way MANOVA^a

1.5 $p = 2, m \geq 2,$

Two-way MANOVA

1.6 $p \geq 3, m \geq 2,$

p -way ANOVA

^aMANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

Two-way ANOVA
Two-way MANOVA
Two-way ANCOVA
Two-way MANCOVA

Indep. v.s. Dependent

1. Categorical v.s. Continuous

1.1 $p = 1, m = 1,$

One-way ANOVA

1.2 $p = 2, m = 1,$

Two-way ANOVA

1.3 $p \geq 3, m = 1,$

p -way ANOVA

1.4 $p = 1, m \geq 2,$

One-way MANOVA^a

1.5 $p = 2, m \geq 2,$

Two-way MANOVA

1.6 $p \geq 3, m \geq 2,$

p -way ANOVA

^aMANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

Indep. v.s. Dependent

1. Categorical v.s. Continuous

1.1 $p = 1, m = 1,$

One-way ANOVA

1.2 $p = 2, m = 1,$

Two-way ANOVA

1.3 $p \geq 3, m = 1,$

p -way ANOVA

1.4 $p = 1, m \geq 2,$

One-way MANOVA^a

1.5 $p = 2, m \geq 2,$

Two-way MANOVA

1.6 $p \geq 3, m \geq 2,$

p -way ANOVA

^aMANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

Indep. v.s. Dependent

1. Categorical v.s. Continuous

1.1 $p = 1, m = 1,$	One-way ANOVA
1.2 $p = 2, m = 1,$	Two-way ANOVA
1.3 $p \geq 3, m = 1,$	p -way ANOVA
1.4 $p = 1, m \geq 2,$	One-way MANOVA ^a
1.5 $p = 2, m \geq 2,$	Two-way MANOVA
1.6 $p \geq 3, m \geq 2,$	p -way ANOVA

^aMANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

Indep. v.s. Dependent

1. Categorical v.s. Continuous

1.1 $p = 1, m = 1,$

One-way ANOVA

1.2 $p = 2, m = 1,$

Two-way ANOVA

1.3 $p \geq 3, m = 1,$

p -way ANOVA

1.4 $p = 1, m \geq 2,$

One-way MANOVA^a

1.5 $p = 2, m \geq 2,$

Two-way MANOVA

1.6 $p \geq 3, m \geq 2,$

p -way ANOVA

^aMANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

Indep. v.s. Dependent

1. Categorical v.s. Continuous

1.1 $p = 1, m = 1,$	One-way ANOVA
1.2 $p = 2, m = 1,$	Two-way ANOVA
1.3 $p \geq 3, m = 1,$	p -way ANOVA
1.4 $p = 1, m \geq 2,$	One-way MANOVA ^a
1.5 $p = 2, m \geq 2,$	Two-way MANOVA
1.6 $p \geq 3, m \geq 2,$	p -way MANOVA

^aMANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

2.1 $m_{ind} = 1, m_{dep} = 1,$	Simple linear regression
2.2 $m_{ind} \geq 2$	Multiple linear regression
2.3 $m_{dep} \geq 2$	Multivariate linear regression

Indep. v.s. Dependent

1. Categorical v.s. Continuous

1.1 $p = 1, m = 1,$	One-way ANOVA
1.2 $p = 2, m = 1,$	Two-way ANOVA
1.3 $p \geq 3, m = 1,$	p -way ANOVA
1.4 $p = 1, m \geq 2,$	One-way MANOVA ^a
1.5 $p = 2, m \geq 2,$	Two-way MANOVA
1.6 $p \geq 3, m \geq 2,$	p -way ANOVA

^aMANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

2.1 $m_{ind} = 1, m_{dep} = 1,$	Simple linear regression
2.2 $m_{ind} \geq 2$	Multiple linear regression
2.3 $m_{dep} \geq 2$	Multivariate linear regression

Indep. v.s. Dependent

1. Categorical v.s. Continuous

1.1 $p = 1, m = 1,$	One-way ANOVA
1.2 $p = 2, m = 1,$	Two-way ANOVA
1.3 $p \geq 3, m = 1,$	p -way ANOVA
1.4 $p = 1, m \geq 2,$	One-way MANOVA ^a
1.5 $p = 2, m \geq 2,$	Two-way MANOVA
1.6 $p \geq 3, m \geq 2,$	p -way ANOVA

^aMANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

2.1 $m_{ind} = 1, m_{dep} = 1,$	Simple linear regression
2.2 $m_{ind} \geq 2$	Multiple linear regression
2.3 $m_{dep} \geq 2$	Multivariate linear regression

Indep. v.s. Dependent

1. Categorical v.s. Continuous

1.1 $p = 1, m = 1,$	One-way ANOVA
1.2 $p = 2, m = 1,$	Two-way ANOVA
1.3 $p \geq 3, m = 1,$	p -way ANOVA
1.4 $p = 1, m \geq 2,$	One-way MANOVA ^a
1.5 $p = 2, m \geq 2,$	Two-way MANOVA
1.6 $p \geq 3, m \geq 2,$	p -way ANOVA

^aMANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

2.1 $m_{ind} = 1, m_{dep} = 1,$	Simple linear regression
2.2 $m_{ind} \geq 2$	Multiple linear regression
2.3 $m_{dep} \geq 2$	Multivariate linear regression

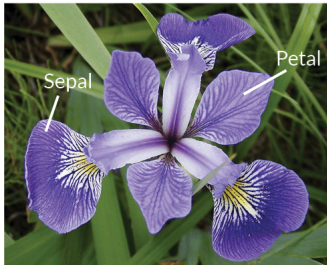
E.g. One example for MANOVA¹.

¹<http://www.sthda.com/english/wiki/manova-test-in-r-multivariate-analysis-of-variance>

E.g. One example for MANOVA¹.



¹<http://www.sthda.com/english/wiki/manova-test-in-r-multivariate-analysis-of-variance>



Iris Versicolor



Iris Setosa



Iris Virginica

```

1 > library(datasets)
2 > data(iris)
3 > summary(iris)
4   Sepal.Length   Sepal.Width   Petal.Length   Petal.Width
5   Species
6   Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100   setosa
7   1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300   versicolor
8   Median :5.800   Median :3.000   Median :4.350   Median :1.300   virginica
9   Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199
10  3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
11  Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500
12 > my_data <- iris
13 > my_data
14   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
15 1           5.1         3.5         1.4         0.2   setosa
16 2           4.9         3.0         1.4         0.2   setosa
17 3           4.7         3.2         1.3         0.2   setosa
18 4           4.6         3.1         1.5         0.2   setosa
19 5           5.0         3.6         1.4         0.2   setosa
20 6           5.4         3.9         1.7         0.4   setosa
21 7           4.6         3.4         1.4         0.3   setosa
22 8           5.0         3.4         1.5         0.2   setosa
23 9           4.4         2.9         1.4         0.2   setosa
24 10          4.9         3.1         1.5         0.1   setosa

```

```

1 > # Compute MAOVA test now
2 > res.man <- manova(cbind(Sepal.Length, Petal.Length) ~ Species, data = iris)
3 > summary(res.man)
4           Df Pillai approx F num Df den Df   Pr(>F)
5 Species      2 0.9885  71.829      4   294 < 2.2e-16 ***
6 Residuals 147
7 ---
8 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
9 > # Look to see which differ
10 > summary.aov(res.man)
11 Response Sepal.Length :
12           Df Sum Sq Mean Sq F value Pr(>F)
13 Species      2 63.212  31.606  119.26 < 2.2e-16 ***
14 Residuals  147 38.956   0.265
15 ---
16 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
17
18 Response Petal.Length :
19           Df Sum Sq Mean Sq F value Pr(>F)
20 Species      2 437.10 218.551 1180.2 < 2.2e-16 ***
21 Residuals  147 27.22   0.185
22 ---
23 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1:w

```

Concl.: Two variables are highly significantly different among species.