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

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

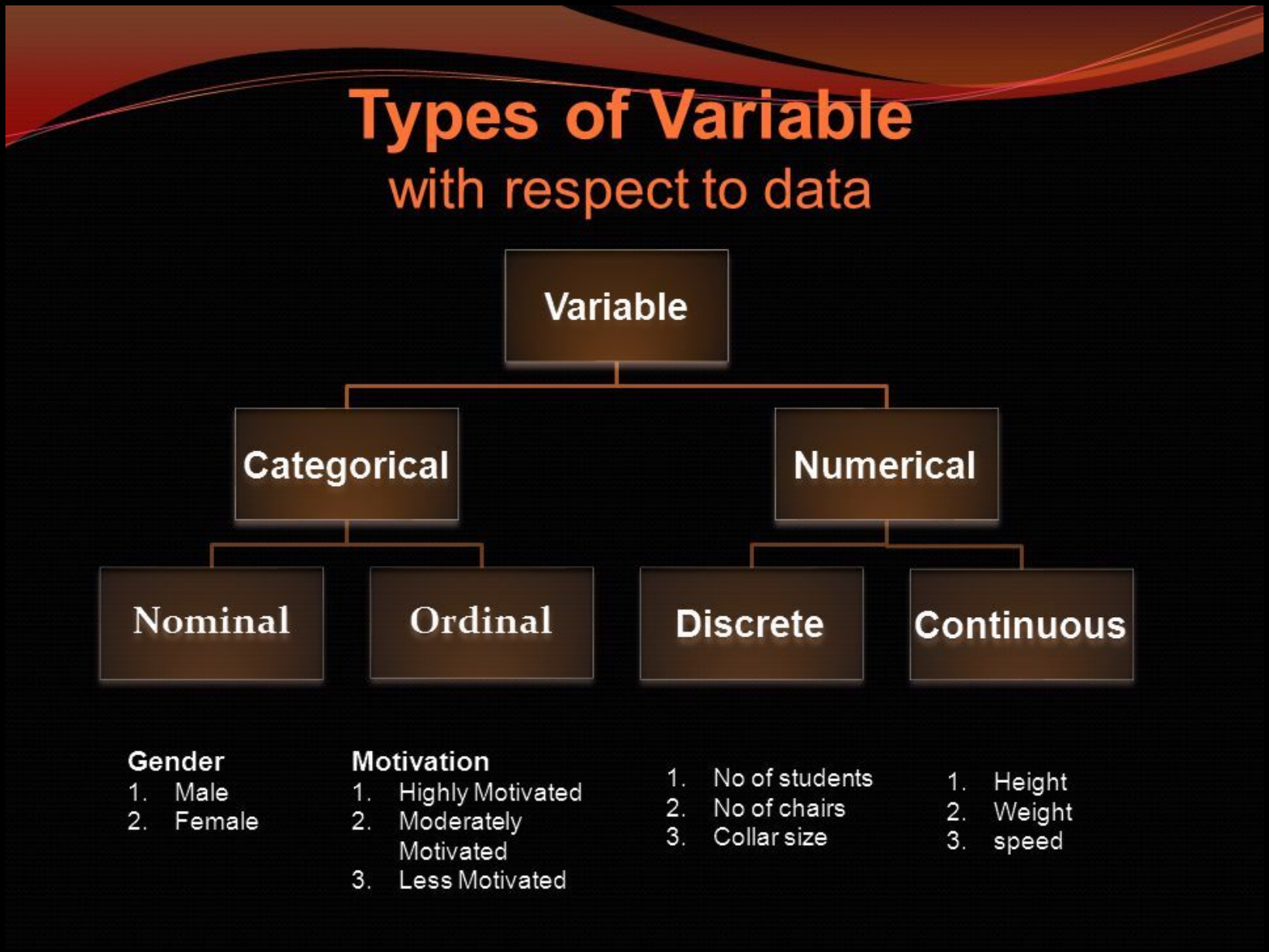
§ 13.1 Introduction

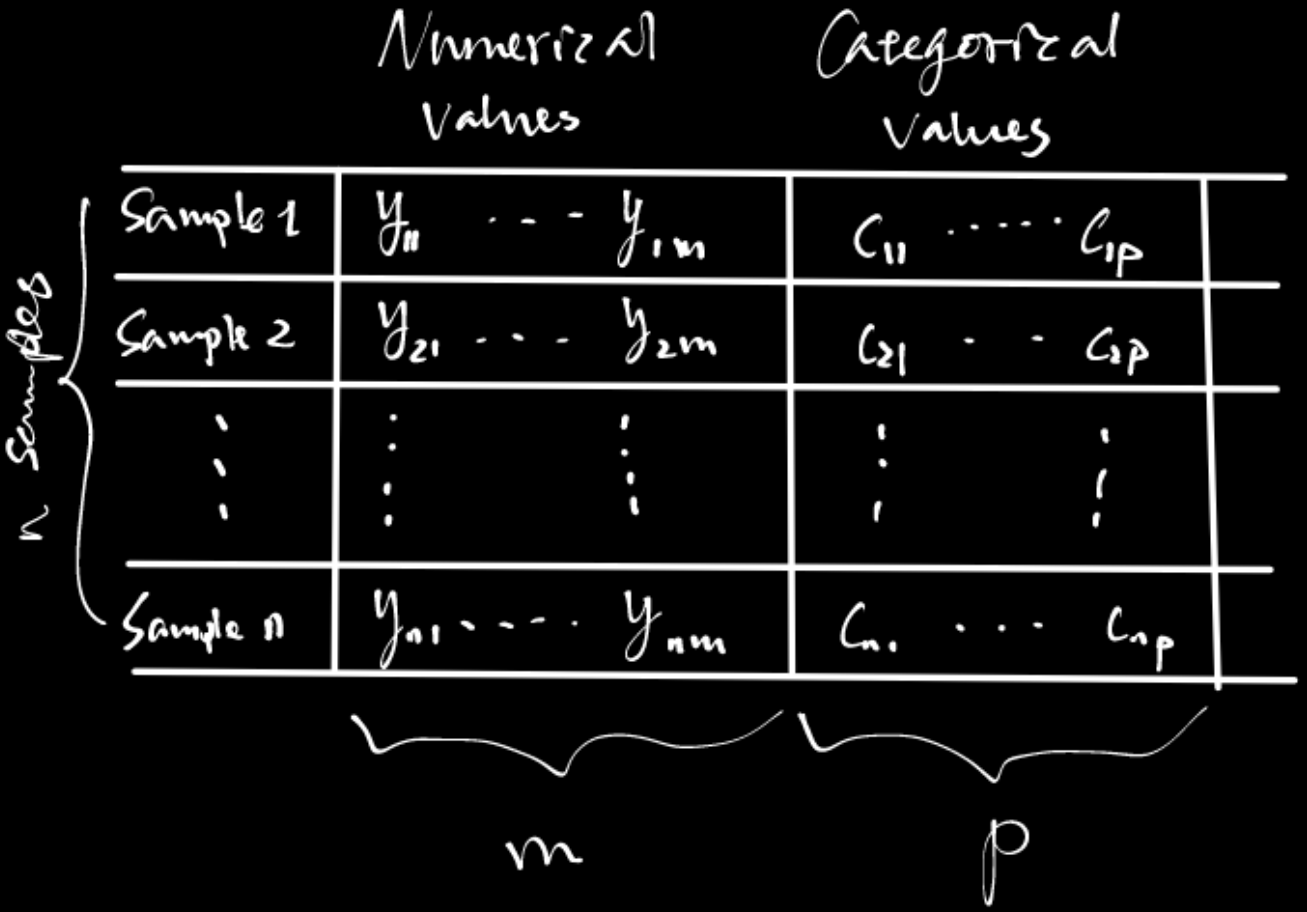
§ 13.2 The *F* Test for a Randomized Block Design

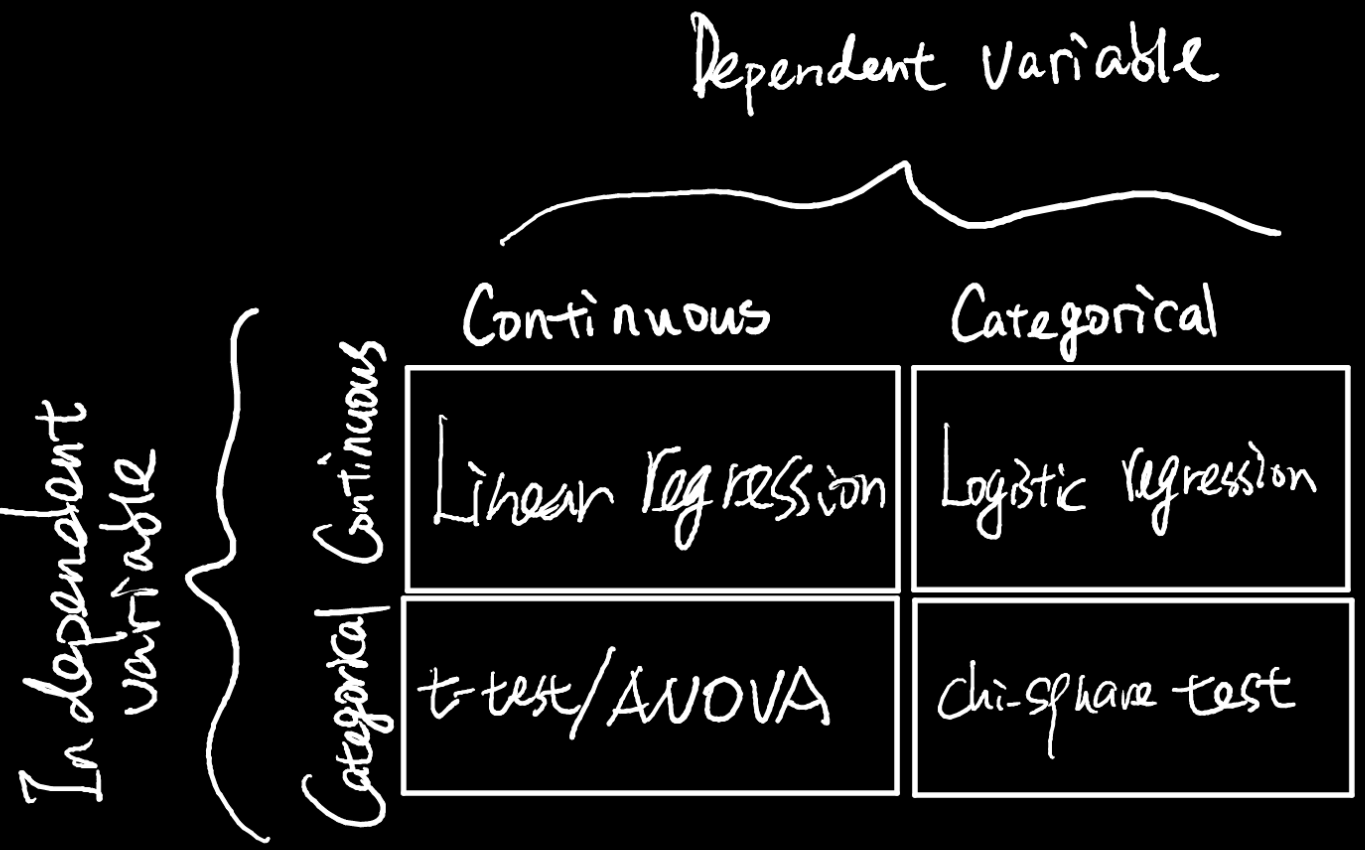
§ 13.A Appendix: Some Discussions and Extensions

Chapter 13. Randomized Block Designs

§ 13.A Appendix: Some Discussions and Extensions




Numerical Values
Categorical Values
n Samples
Sample 1
Sample 2
Sample n
y_{11} ... y_{1m}
y_{21} ... y_{2m}
y_{n1} ... y_{nm}
c_{11} ... c_{1p}
c_{21} ... c_{2p}
c_{n1} ... c_{np}
m
p


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

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 $p >= 3, m == 1$ → p-way ANOVA
   - 1.4 $p == 1, m >= 2$ → One-way MANOVA[2]
   - 1.5 $p == 2, m >= 2$ → Two-way MANOVA
   - 1.6 $p >= 3, m >= 2$ → p-way ANOVA

[2] MANOVA refers to the multivariate analysis of variance. ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous
   - 2.1 $m_{dep} == 1, m_{indep} == 1$ → Simple linear regression
   - 2.2 $m_{dep} == 1, m_{indep} >= 2$ → Multiple linear regression
   - 2.3 $m_{dep} >= 2$ → 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 $p > 2$, $m == 1$. — p-way ANOVA
   - 1.4 $p == 1$, $m \geq 2$. — One-way MANOVA[3]
   - 1.5 $p == 2$, $m \geq 2$. — Two-way MANOVA
   - 1.6 $p > 2$, $m \geq 2$. — p-way ANOVA

[3] MANOVA refers to the multivariate analysis of variance. ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous
   - 2.1 $m_{IS} == 1$, $m_{DS} == 1$ — Simple linear regression
   - 2.2 $m_{IS} > 1$ — Multiple linear regression
   - 2.3 $m_{DS} > 1$ — 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

[a] MANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

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

[a] MANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

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

[a] MANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

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

[a] MANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

2.1	$p = 1, m = 1$	Simple linear regression
2.2	$p \geq 2, m = 1$	Multiple linear regression
2.3	$p \geq 2, m \geq 2$	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

[a] MANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

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

[a] MANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

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

[a] MANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

2.1	$p = 1, m = 1$	Simple linear regression
2.2	$p \geq 2, m = 1$	Multiple linear regression
2.3	$p \geq 2, m \geq 2$	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

[a] MANOVA refers to the multivariate analysis of variance
ANOVA refers to the univariate analysis of variance.

2. Continuous v.s. Continuous

2.1	$p = 1, m = 1$	Simple linear regression
2.2	$p \geq 2, m = 1$	Multiple linear regression
2.3	$p \geq 2, m \geq 2$	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

[a] MANOVA 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

[a] MANOVA 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

[a] MANOVA 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

[a] MANOVA 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 MANOVA[1].

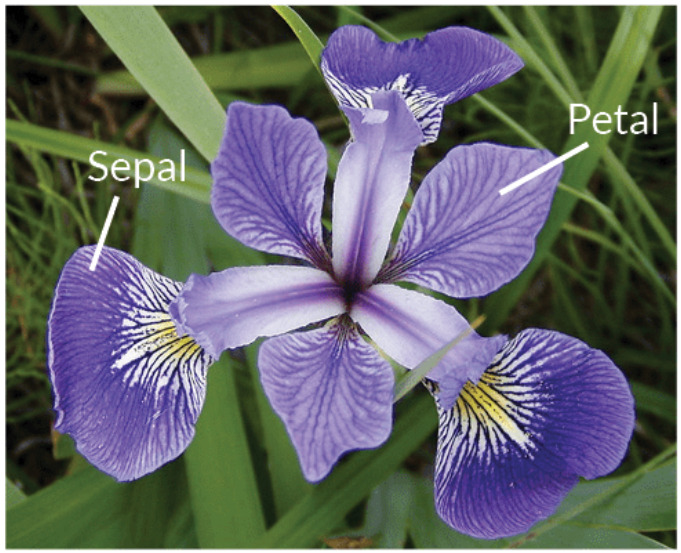
[1] http://www.sthda.com/english/wiki/manova-test-in-r-multivariate-analysis-of-variance

E.g. One example for MANOVA[1].

The Iris Dataset

Collected by Ronald Fisher in 1936

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



Iris Versicolor

Iris Setosa

Iris Virginica

```
> library (datasets)
> data( iris )
> summary(iris)
   Sepal.Length    Sepal.Width     Petal.Length    Petal.Width
          Species
 Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100   setosa
         :50
 1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300   versicolor
         :50
 Median :5.800   Median :3.000   Median :4.350   Median :1.300   virginica
         :50
 Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199
 3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
 Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500
> my_data <- iris
> my_data
    Sepal.Length Sepal.Width Petal.Length Petal.Width    Species
1            5.1         3.5          1.4         0.2     setosa
2            4.9         3.0          1.4         0.2     setosa
3            4.7         3.2          1.3         0.2     setosa
4            4.6         3.1          1.5         0.2     setosa
5            5.0         3.6          1.4         0.2     setosa
6            5.4         3.9          1.7         0.4     setosa
7            4.6         3.4          1.4         0.3     setosa
8            5.0         3.4          1.5         0.2     setosa
9            4.4         2.9          1.4         0.2     setosa
10           4.9         3.1          1.5         0.1     setosa
```

```
> # Compute MAOVA test now
> res.man <- manova(cbind(Sepal.Length, Petal.Length) ~ Species, data = iris)
> summary(res.man)
             Df  Pillai  approx F num Df den Df    Pr(>F)
Species       2 0.9885   71.829      4     294 < 2.2e-16 ***
Residuals   147
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> # Look to see which differ
> summary.aov(res.man)
 Response Sepal.Length :
              Df Sum Sq Mean Sq F value Pr(>F)
Species        2 63.212  31.606  119.26 < 2.2e-16 ***
Residuals    147 38.956   0.265
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

 Response Petal.Length :
              Df Sum Sq Mean Sq F value Pr(>F)
Species        2 437.10 218.551 1180.2 < 2.2e-16 ***
Residuals    147  27.22   0.185
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1:w
```

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