Analysis of Variance (ANOVA)

Matt Arthur

UCR GradQuant

18 Jan 2022

Overview

1. Introduction to ANOVA
2. One-Way ANOVA
   - Example
   - One-Way ANOVA Table
3. ANOVA Assumptions
   - Normality
   - Constant Variance
   - Remedial Measures
4. Two-Way ANOVA
5. Post Hoc Comparisons
6. Random Effects

Overview

1. Introduction to ANOVA
2. One-Way ANOVA
   - Example
   - One-Way ANOVA Table
3. ANOVA Assumptions
   - Normality
   - Constant Variance
   - Remedial Measures
4. Two-Way ANOVA
5. Post Hoc Comparisons
6. Random Effects

Introduction

What is ANOVA?

- Analysis of Variance (ANOVA) is a collection of statistical models and their associated estimation procedures used to analyze differences among group means.

Introduction

What is ANOVA?

- Analysis of Variance (ANOVA) is a collection of statistical models and their associated estimated procedures used to analyze differences among group means.
- It's not called *Analysis of Means* because we make inferences about means by analyzing variance.

Introduction

What is ANOVA?

- Analysis of Variance (ANOVA) is a collection of statistical models and their associated estimated procedures used to analyze differences among group means.
- It's not called *Analysis of Means* because we make inferences about means by analyzing variance.
- Assumptions:
  - Data must be *Normally Distributed*
  - Data must be *Independent*
  - Data must have *Equal Population Variances*

Introduction

ANOVA Types:

- Various types of ANOVA reflect different experimental designs and situations for which they have been developed.
- *n*-way ANOVA refers to *n Independent Variables* used in ANOVA tests.

Introduction

Examples in which ANOVA could be applied:

- A group of patients is trying three different treatments. You want to see if one if better than the others.
- A manufacturer has two different processes to make light bulbs. They want to know if one process is better than the other
- Students from different colleges take the same exam. You want to see if one college outperforms the other.

Overview

1. Introduction to ANOVA
2. One-Way ANOVA
   - Example
   - One-Way ANOVA Table
3. ANOVA Assumptions
   - Normality
   - Constant Variance
   - Remedial Measures
4. Two-Way ANOVA
5. Post Hoc Comparisons
6. Random Effects

One-Way ANOVA

One-Way ANOVA refers to the situation in which we want to compare k means corresponding to different levels of a single independent variable.

General Setup:

$$\begin{aligned} x_{11}, x_{12}, \ldots, x_{1n_1} &\sim \mathrm{N}\left(\mu_1, \sigma^2\right) \\ x_{21}, x_{22}, \ldots, x_{2n_2} &\sim \mathrm{N}\left(\mu_2, \sigma^2\right) \\ &\vdots \\ x_{k1}, x_{k2}, \ldots, x_{kn_k} &\sim \mathrm{N}\left(\mu_k, \sigma^2\right) \end{aligned}$$

One-Way ANOVA Notation

From the previous slide, our assumption is:

$$X_{ij} \sim \mathrm{N}\left(\mu_i, \sigma^2\right)$$

For $i = 1, \ldots, k$, $j = 1, \ldots, n_i$

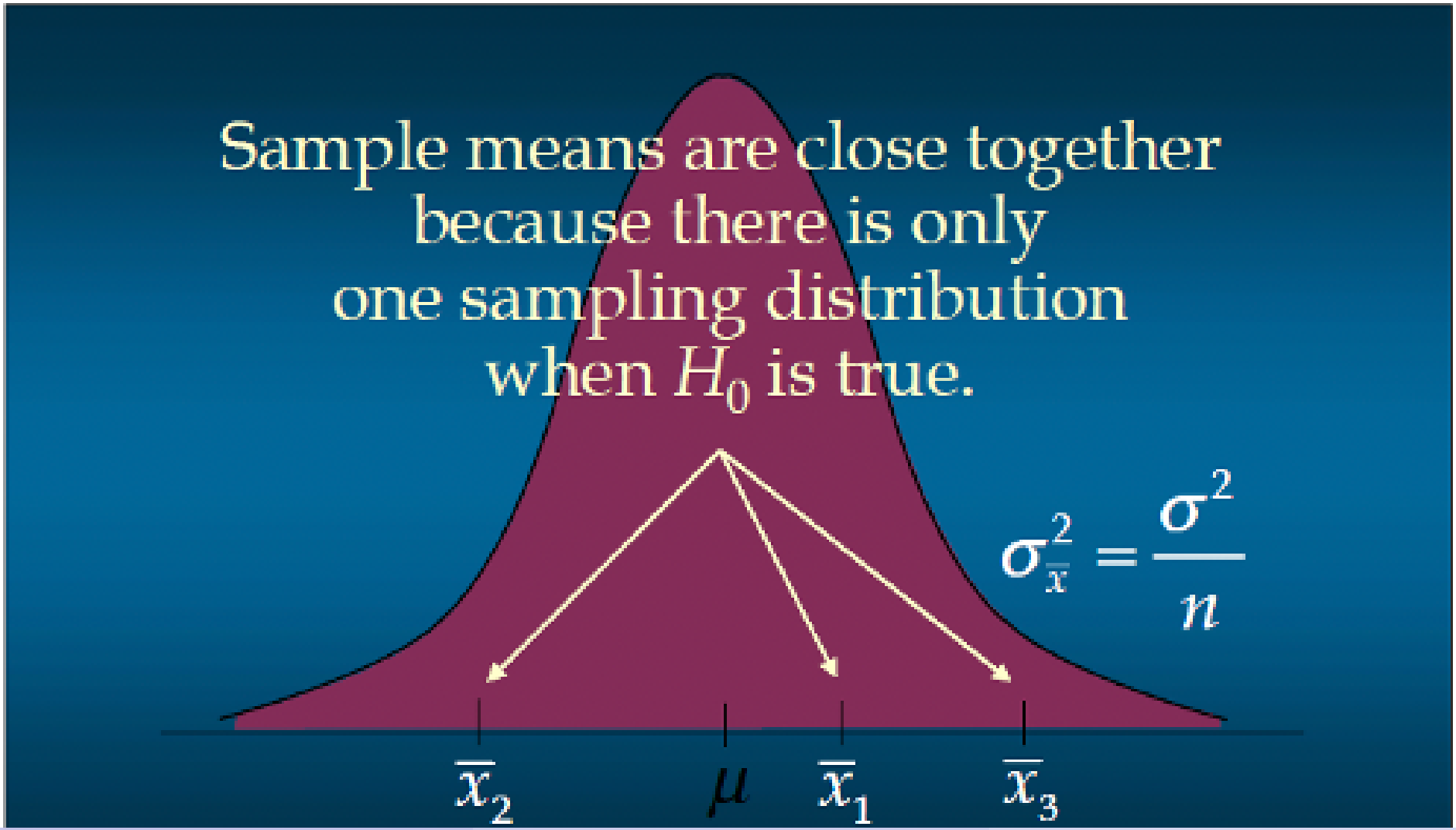
Visual Interpretation

We want to use results from our sample to test the following hypothesis:

- $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$
- H_1 : at least one μ_i differs from the others, for $i = 1, \ldots, k$
- If H_0 is rejected, we cannot conclude that all population means are the same.
- Rejecting the null hypothesis means that at least two of the group population means are different from one another.
- If we fail to reject H_0, it is reasonable to assume that all population means are the same.

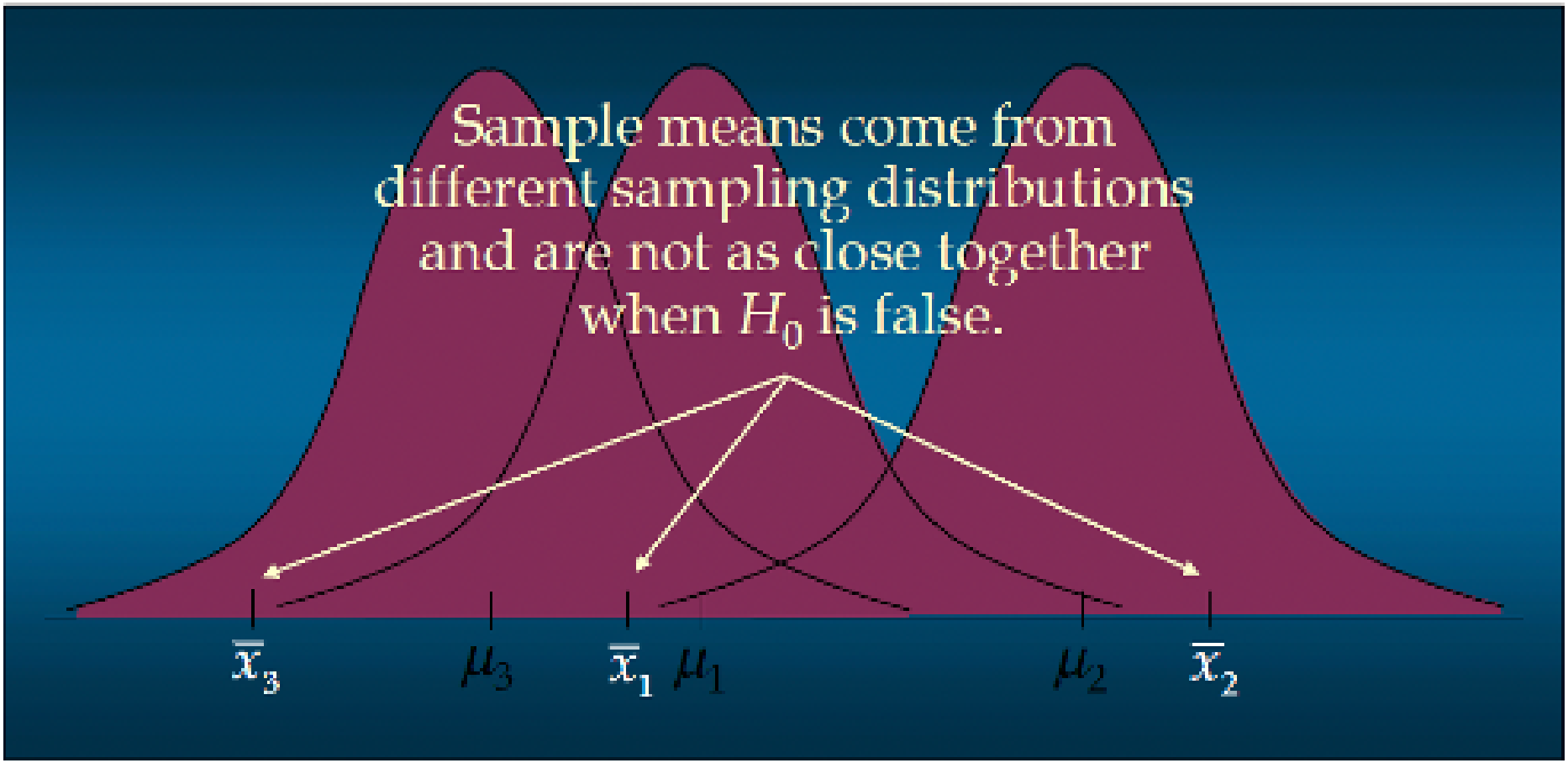
Visual Interpretation

Sample distribution of $\bar{X}$ given H_0 is true:



Visual Interpretation

Sample distribution of $\bar{X}$ given H_0 is false (at least two means differ)



One-Way ANOVA Notation

From the previous slide, our assumption is:

$$X_{ij} \sim \mathrm{N}\left(\mu_i, \sigma^2\right)$$

For $i = 1, \ldots, k$, $j = 1, \ldots, n_i$

Let $n = n_1 + n_2 + \cdots + n_k$

One-Way ANOVA Notation

Let $\bar{X}_{i\cdot}$ denote the *within-group* mean for group i:

$$\bar{x}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$$

One-Way ANOVA Notation

Let $\bar{X}_{i\cdot}$ denote the *within-group* mean for group i:

$$\bar{x}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$$

Let $\bar{X}_{\cdot\cdot}$ denote the *grand* mean:

$$\bar{x}_{\cdot\cdot} = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}$$

One-Way ANOVA Notation

The idea with ANOVA is to compare the variability between groups and within groups. To measure between-group variability, we use a between-group sum of squares, SSTR:

$$\text{SSTR} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot}\right)^2$$

To measure the within-group variability, we use an error sum of squares:

$$\text{SSE} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(X_{ij} - \bar{X}_{i\cdot}\right)^2$$

One-Way ANOVA Notation

To measure between-group variability, we use a between-group sum of squares:

$$\text{SSTR} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot}\right)^2 = \sum_{i=1}^{k} n_i \left(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot}\right)^2$$

To measure the within-group variability, we use an error sum of squares:

$$\text{SSE} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(X_{ij} - \bar{X}_{i\cdot}\right)^2 = \sum_{i=1}^{k} \left(n_i - 1\right) S_i^2$$

To measure the total variability of the observed data, use a total sum of squares:

$$\text{SST} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(X_{ij} - \bar{X}_{\cdot\cdot}\right)^2$$

One-Way ANOVA Equations

S_i^2 is the sample variance within group i:

$$S_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} \left(X_{ij} - \bar{X}_{i\cdot} \right)^2$$

One-Way ANOVA Equations

SSTR and SSE both capture the variability in the observations about their sample means. Both can be used to estimate σ^2:

- Define $\text{MSE} = \text{SSE}/(n-k)$
- Define $\text{MSTR} = \text{SSTR}/(k-1)$

One-Way ANOVA Equations

SSTR and SSE both capture the variability in the observations about their sample means. Both can be used to estimate σ^2:

- Define $\text{MSE} = \text{SSE}/(n-k)$
- Define $\text{MSTR} = \text{SSTR}/(k-1)$

Equations

- MSTR only estimates σ^2 if the population means are equal
- If the population means are not equal, MSTR estimates a quantity larger than σ^2
- If the null hypothesis is true and the ANOVA assumptions are valid, the sampling distribution of

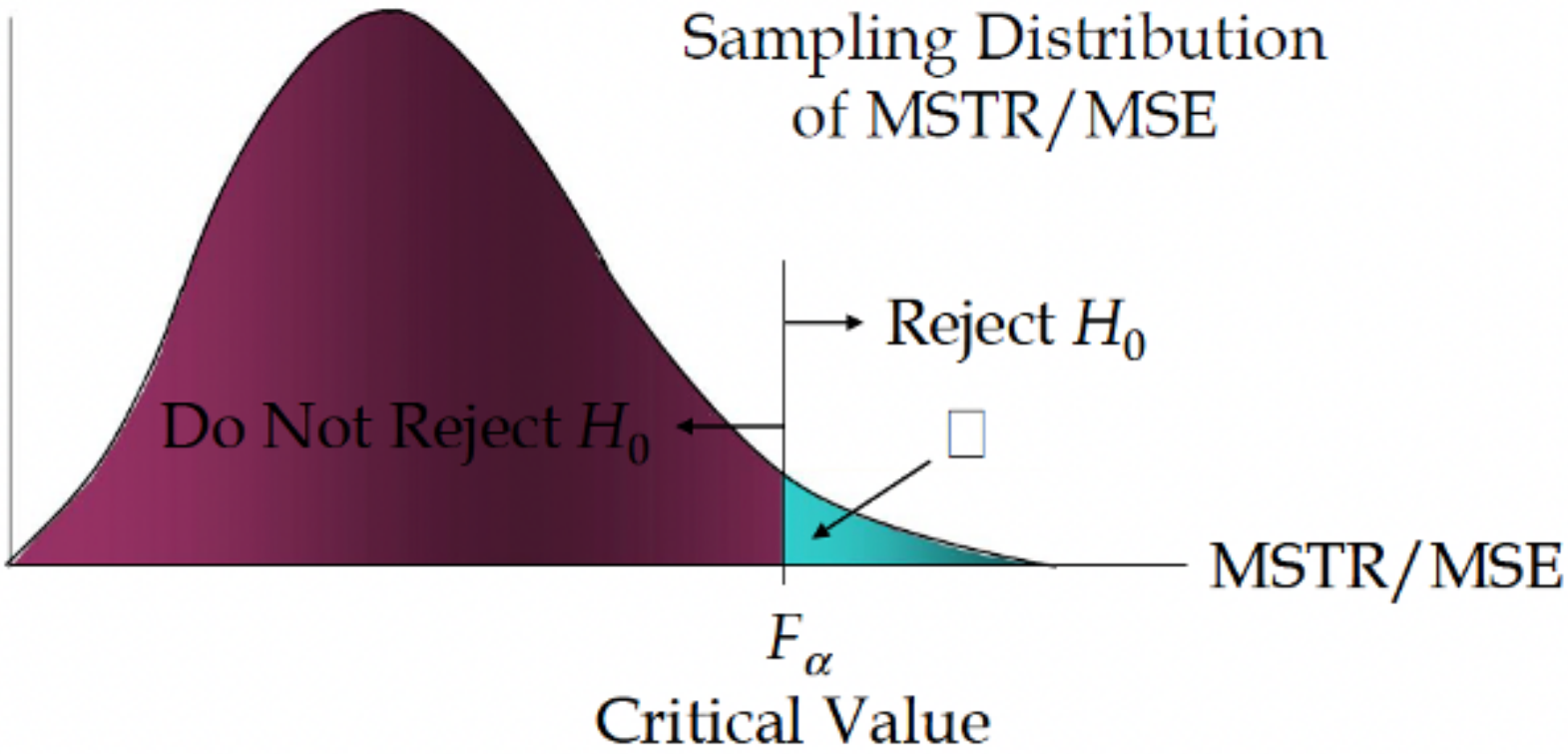
$$F = \frac{\text{MSTR}}{\text{MSE}}$$

follows an F-distribution with degrees of freedom $k - 1$ and $n - k$

Equations

- If the meas of the k populations are not equal, the value of F will be inflated because MSTR overestimates σ^2.
- Thus, we will reject the null hypothesis if the resulting value of MSTR/MSE appears too large to have been selected at random from the appropriate F-distribution.

Equations


Sampling Distribution
of MSTR/MSE
Reject H_0
Do Not Reject H_0
MSTR/MSE
F_α
Critical Value

Example

A pharmaceutical company conducts an experiment to test the effect of a new cholesterol medication. The company selection 15 subjects randomly from a larger population. Each subject is randomly assigned to one of three treatment groups. Within each treatment group, subjects receive a different dose of the new medication.

- In Group 1, subjects receive 0mg/day
- In Group 2, subjects receive 50mg/day
- In Group 3, subjects receive 100mg/day

Example

- In Group 1, subjects receive 0mg/day
- In Group 2, subjects receive 50mg/day
- In Group 3, subjects receive 100mg/day

After 30 days, doctors measure the cholesterol level of each subject. The results for all 15 subjects appear in the following table.

Example

Group 1 (0mg)	Group 2 (50mg)	Group 3 (100mg)
210	210	180
240	240	210
270	240	210
270	270	210
300	270	240

Interest: Does dosage level have a significant effect on cholesterol level?

Example

First, calculate the sample mean, grand mean, and sample variance for each group.

- Sample Means:

	Group 1 (0mg)	Group 2 (50gm)	Group 3 (100mg)
Mean	258	246	210

- Sample Variances:

	Group 1 (0mg)	Group 2 (50mg)	Group 3 (100mg)
Variance	1170	630	450

- Grand Mean: $(258 + 256 + 210)/3 = 238$

Example

Then, calculate the within- and between-group estimate of variance:

$$\text{MSE} = \frac{(5-1)*1170 + (5-1)*630 + (5-1)*450}{15-3} = \frac{9000}{12} = 750$$

$$\text{MSTR} = \frac{5*(258-238)^2 + 5*(246-238)^2 + 5*(210-238)^2}{3-1} = 3120$$

The observed F-value is:

$$F = \frac{\text{MSTR}}{\text{MSE}} = \frac{3120}{750} = 4.16$$

Example

- Under the null hypothesis, the observed F
- Using an F-table, the critical value for an F-distribution with degrees of freedom 2 and 12 is 3.8853 (significance level $= 0.05$)
- Since $F = 4.16 > 3.8853$, we reject the null hypothesis and conclude that the mean cholesterol levels are **not** the same among the three treatments.

ANOVA Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F
Treatment	SSTR	$k-1$	MSTR	MSTR/MSE
Error	SSE	$n-k$	MSE	
Total	SST	$n-1$		

ANOVA Table

The total sum of squares (SST) measures the total vartiability in the response variance, without considering the group structure:

$$\text{SST} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(X_{ij} - \bar{X}_{..}\right)^2$$

ANOVA can be viewed as the process of partitioning the total sum of squares into their corresponding components: Treatments and Error

ANOVA Table

$$(X_{ij} - \bar{X}_{..}) = (X_{ij} - \bar{X}_{i.}) + (X_{i.} - \bar{X}_{..})$$

This relationship also holds if we square and sum all terms:

$$\underbrace{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_{..})^2}_{\text{SST}} = \underbrace{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_{i.})^2}_{\text{SSE}} + \underbrace{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{i.} - \bar{X}_{..})^2}_{\text{SSTR}}$$

ANOVA Table

R can also perform one-way ANOVA:

```
doseC <- c(210,240,270,270,300,210,240,240,270,270,180,210,
           210,210,240)
doseG <- c(rep("group1",5), rep("group2",5),rep("group3",5))
dose <- data.frame(dose=as.numeric(doseC),
                   group=factor(doseG))
head(dose)

##   dose  group
## 1  210 group1
## 2  240 group1
## 3  270 group1
## 4  270 group1
## 5  300 group1
## 6  210 group2
```

ANOVA Table

```
summary(aov(dose$dose ~ dose$group))

##              Df Sum Sq Mean Sq F value Pr(>F)
## dose$group    2   6240    3120    4.16 0.0424 *
## Residuals    12   9000     750
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

Overview

1. Introduction to ANOVA
2. One-Way ANOVA
   - Example
   - One-Way ANOVA Table
3. ANOVA Assumptions
   - Normality
   - Constant Variance
   - Remedial Measures
4. Two-Way ANOVA
5. Post Hoc Comparisons
6. Random Effects

ANOVA Assumptions

To conduct ANOVA tests, we need to check model assumptions. The ANOVA model is

$$X_{ij} = \mu_i + \epsilon_{ij}, \quad \epsilon_{ij} \sim \mathrm{N}\left(0, \sigma^2\right)$$

where ϵ_{ij} are independent. Thus, we have three assumptions to make:

1. Normality
2. Equal, Constant Variance
3. Independence

We can check these assumptions using our residuals

ANOVA Assumptions

Residuals are calculated as

$$\hat{\epsilon}_{ij} = X_{ij} - \hat{\mu}_i = X_{ij} - \bar{X}_{i\cdot}$$

In our Example:

Group 1 (residuals)	**Group 2 (residuals)**	**Group 3 (residuals)**
210 (210 - 258 = -48)	210 (210 - 246 = -36)	180 (180 - 210 = -30)
240 (240 - 258 = -18)	240 (240 - 246 = -6)	210 (210 - 210 = 0)
270 (270 - 258 = 12)	240 (240 - 246 = -6)	210 (210 - 210 = 0)
270 (270 - 258 = 12)	270 (270 - 246 = 24)	210 (210 - 210 = 0)
300 (300 - 258 = 42)	270 (270 - 246 = 24)	240 (240 - 210 = 30)

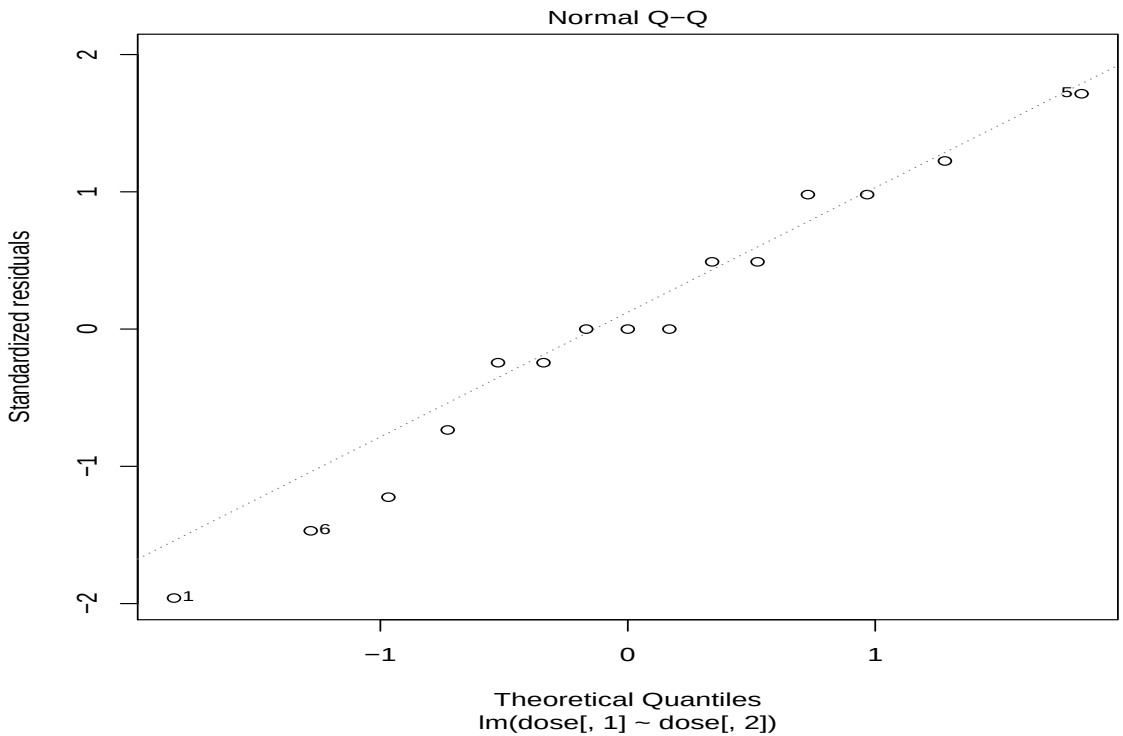
Normality Assumption

To check the normality assumption, we can use a Q-Q Plot:

```
residuals <- c(-48,-18,12,12,42,-36,-6,-6,24,24-30,0,0,0,
               30)
m <- lm(dose[,1] ~ dose[,2])
```

Normality

```
plot(m, which = 2)
```



Normality

Points in the Q-Q Plot should form an approximate 45-degree line:

- Indicating that the theoretical quantiles from a normal distribution match the estimated quantiles from the data.

Normality

We an also test for normality using the Shapiro-Wilk Normality Test:

```
shapiro.test(residuals)

##
##  Shapiro-Wilk normality test
##
## data:  residuals
## W = 0.96461, p-value = 0.7976
```

In this case, the null hypothesis is:

$$H_0 : \text{Data are drawn from a Normal Distribution}$$

Normality

According to both the Q-Q Plot and the Shapiro-Wilk test, we can conclude that our data are normality distributed

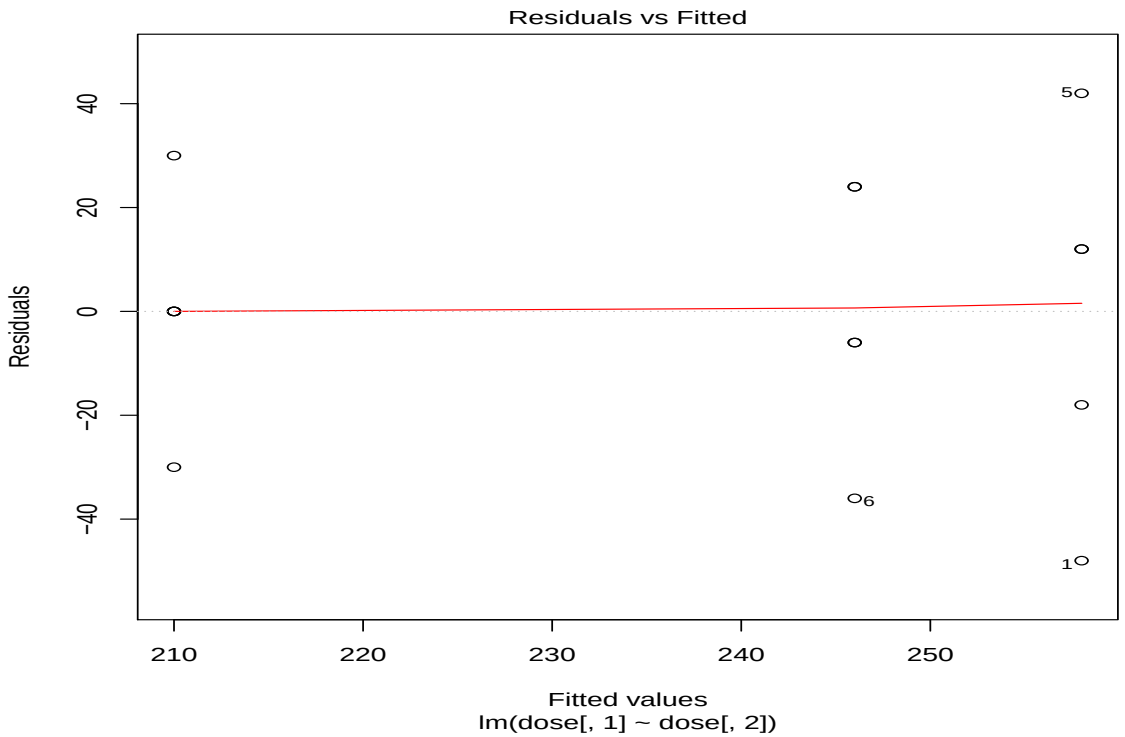
Constant Variance of Errors

There are also a few ways to check for homogeneity of variances. The easiest way is to use a residual plot.

If the variances are equal, the residuals should be distributed nearly equally above and below zero.

Constant Variance

```
plot(m, which = 1)
```



Constant Variance

Another way to test for homogeneity of variance is to use Bartlett's Test, which tests the hypothesis:

$$H_0 : \text{All Variances are Equal}$$

```
bartlett.test(x = dose[,1], g = dose[,2])

##
##  Bartlett test of homogeneity of variances
##
## data:  dose[, 1] and dose[, 2]
## Bartlett's K-squared = 0.86578, df = 2, p-value = 0.6486
```

Constant Variance

Since the p-value is large, we fail to reject H_0 and conclude that all variances are equal.

Remedial Measures

If the assumptions about normality, constant variance, and independence are not satisfied, our ANOVA results may be unreliable.

- To remedy these issues, we can attempt to transform the data.

Remedial Measures

To remedy these issues, we can attempt to transform the data.

- Square Root Transformation.

$$X_{ij}^{\star} = \sqrt{X_{ij}}$$

  Useful if σ^2 is proportional to μ_i. I.e., $\sigma^2 = \sigma_i^2 = c\mu_i$ for some c.

- Logarithmic Transformation.

$$X_{ij}^{\star} = \log(X_{ij})$$

  Useful if the standard deviation (σ) is proportional to μ_i

- Inverse Transformation:

$$X_{ij}^{\star} = \frac{1}{X_{ij}}$$

Remedial Measures

To remedy these issues, we can attempt to transform the data.

- Box-Cox transformation:

$$X_{ij}^{\star}(\lambda) = \begin{cases} \frac{X_{ij}^{\lambda}-1}{\lambda} & \text{if } \lambda \neq 0 \\ \log(X_{ij}) & \text{if } \lambda = 0 \end{cases}$$

These transformations can help to address non-constant variance. In some cases, they may also address non-normal data.

- They will not address dependent data, and may not be effective in all situations.
- They may also have an effect of the interpretability of the model.

Overview

1. Introduction to ANOVA
2. One-Way ANOVA
   - Example
   - One-Way ANOVA Table
3. ANOVA Assumptions
   - Normality
   - Constant Variance
   - Remedial Measures
4. Two-Way ANOVA
5. Post Hoc Comparisons
6. Random Effects

Introduction to Two-Way ANOVA

Until now, we have dealt only with the case where we have one treatment.

- What if we have multiple grouping variables?
- What if we are using repeated-measures data?
- What if we want to include a covariate?

Two-Way ANOVA allows us to add a second grouping variable. The basic ideas are the same, but the math is a bit more complicated.

Notation

When we have two categorical explanatory variables (Factor A, and Factor B):

- Continuous Response Variable: Y_{ijk} = outcome for the kth replication of the ith level of Factor A and the jth level of Factor B
- Factor A has a levels: $i = 1, \ldots, a$
- Factor B has b levels: $j = 1, \ldots, b$
- Observations in cell (i, j) are indexed by k, where $k = 1, 2, \ldots, n_{ij}$
- We say the design is *balanced* if $n_{ij} = n$ for all i, j

Example

Randomly Selected 36 People from a Car Market that bought the same model of car. Suppose these people received different cash offers from their dealer. The 36 customers represent 3 age-based categories: Young, Middle-Aged, Elderly. There are 12 customers in each age-category. Moreover, half of the 36 customers identified as Male, the other half Female.

Our Interest: Does Age or Gender of customers affect the amount of cash offered?

Example

Suppose we randomly selected 36 People from a Car Market who all bought the same model of car. Suppose these people received different cash offers from their dealer. The 36 customers represent 3 age-based categories: Young, Middle-Aged, Elderly. There are 12 customers in each age-category. Moreover, half of the 36 customers identified as Male, the other half Female.

Our Interest: Does age or sex of customers affect the amount of cash offered?

- Now, we are interested in two factors. Age (3 levels) and Sex (2 levels).
- The total number of observations is 36. There are 6 observations per Age/Sex combination

Example

The cash offer results are given in the following table (unit = thousands):

Group		Age					
		Young		Middle		Elderly	
Sex	Female	21	22	30	28	25	21
		23	22	29	27	22	22
		19	23	26	27	23	21
	Male	21	21	26	28	23	21
		22	19	29	27	19	20
		20	25	27	29	20	20

Example

We write a fixed linear model to explain the data:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

where $\epsilon_{ijk} \sim \mathrm{N}\left(0, \sigma^2\right)$, and are independent.

- Grand Mean (μ):

$$\hat{\mu} = \bar{Y}_{...}$$

- Main Effects (α_i, β_j):

$$\alpha_i \to \bar{Y}_{i..} - \bar{Y}_{...}, \quad \beta_j \to \bar{Y}_{.j.} - \bar{Y}_{...}$$

- Interaction Term: $(\alpha\beta)_{ij}$

$$(\alpha\beta)_{ij} \to \bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}$$

Example

Again, for Two-Way ANOVA, we partition variability into components:

$$\begin{aligned} Y_{ijk} - Y_{...} = & \quad (Y_{i..} - Y_{...}) \\ & + \quad (Y_{.j.} - Y_{...}) \\ & + \quad (Y_{ijk} - Y_{ij.}) \\ & + \quad (Y_{ij.} - Y_{i..} - Y_{.j.} + Y_{...}) \end{aligned}$$

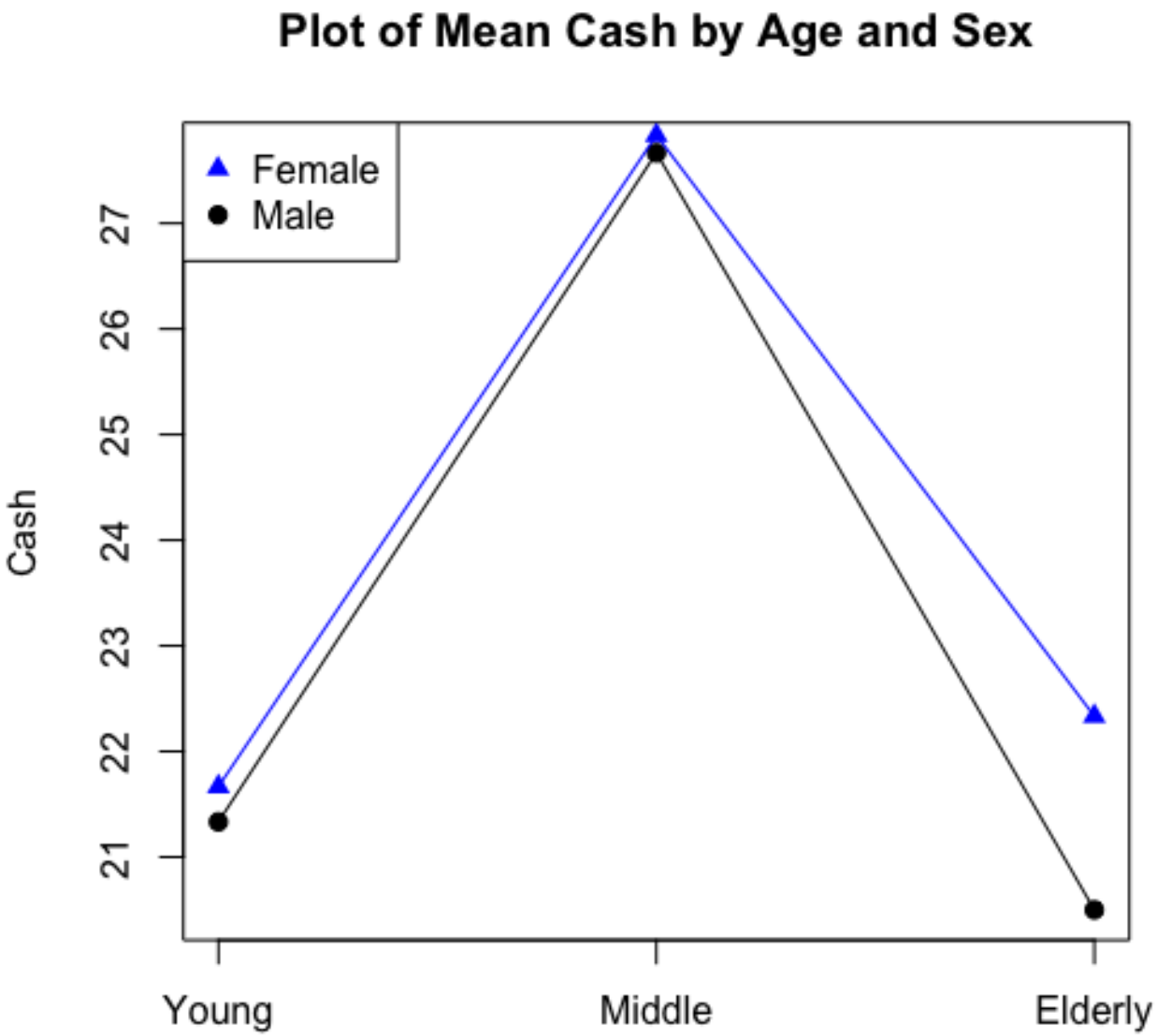
And this relationship again holds if we square and sum over all i, j, k.

Example

Why do we introduce an interaction term?

- For example: If you ask a car dealer, "Do you prefer giving more cash to Males vs. Females?", and they reply: "It depends how old they are."
- This means there exists an interaction effect.
- When an interaction effect is present, the impact of one factor on the response depends on the level of the other factor.

Example


Plot of Mean Cash by Age and Sex
Female
Male
Cash
21
22
23
24
25
26
27
Young
Middle
Elderly

Example

The cash offer results are given in the following table (unit = thousands):

Group		Age						Mean
		Young		Middle		Elderly		
Sex	Female	21	22	30	28	25	21	$Y_{.1.} = 23.94$
		23	22	29	27	22	22	
		19	23	26	27	23	21	
	Male	21	21	26	28	23	21	$Y_{.2.} = 23.17$
		22	19	29	27	19	20	
		20	25	27	29	20	20	
Mean		$Y_{1..} = 21.5$		$Y_{2..} = 27.75$		$Y_{3..} = 21.42$		$Y_{...} = 23.56$

Example

Two-Way ANOVA allows us to test the following hypotheses:

- Main Effect of Factor A (Age):

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0 \quad \text{vs.} \quad H_1 : \text{at least one level is significant}$$

- Main Effect of Factor B (Sex):

$$H_0 : \beta_1 = \beta_2 = 0 \quad \text{vs.} \quad H_1 : \text{at least one level is significant}$$

- Interaction Effect:

$$H_0 : (\alpha\beta)_{ij} = 0, \text{for all } i, j$$

$$\text{vs.} \quad H_1 : \text{at least one interaction term is significant}$$

Example

For the Two-Way ANOVA Table, we partition the total variability in the observed data into sums of squares for each Factor, and for the interaction:

- $\text{SST} = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (Y_{ijk} - \bar{Y}_{...})^2 = \text{SSA} + \text{SSB} + \text{SSAB} + \text{SSE}$
- The corresponding *degrees of freedom* associated with SSA, SSB, and SSAB are $a-1$, $b-1$, and $(a-1)(b-1)$, respectively.
- Decisions about the preceding hypotheses are made by comparing ratios of mean squares. For instance, we could make a statement about Factor A by using the test statistic:

$$F_A = \frac{\text{SSA}/(a-1)}{\text{SSE}/(ab(n-1))}$$

Example

While it is definitely possible to conduct Two-Way ANOVA by hand, I do not recommend doing so. Any statistical software should be able to do this for you. In R, we can conduct a Two-Way ANOVA as follows.

Example

First, read the data in long form:

```
cash <- c(21,22,23,22,19,23,21,21,22,19,20,25,30,28,29,27,
          26,27,26,28,29,27,27,29,25,21,22,22,23,21,23,21,
          19,20,20,20)
age <- factor(c(rep("young", 12), rep("middle", 12),
                rep("elderly", 12)))
sex <- rep(c(rep("female", 6), rep("male", 6)),3)
frame <- data.frame(cash, age, sex)
```

Example

```
head(frame, 8)
```

```
##   cash   age    sex
## 1   21 young female
## 2   22 young female
## 3   23 young female
## 4   22 young female
## 5   19 young female
## 6   23 young female
## 7   21 young   male
## 8   21 young   male
```

Example

Next, use the aov() function in almost exactly the same way as before:

```
summary(aov(frame$cash ~ frame$age*frame$sex))
```

```
##                      Df Sum Sq Mean Sq F value   Pr(>F)
## frame$age             2  316.7  158.36  66.291 9.79e-12 ***
## frame$sex             1    5.4    5.44   2.279    0.142
## frame$age:frame$sex   2    5.1    2.53   1.058    0.360
## Residuals            30   71.7    2.39
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

Example

The results of the ANOVA table suggest:

- There is no interaction between sex and age
- Sex does not have an appreciable impact on cash offerings.
- Age, however, does seem to have a significant effect on the amount of cash offered at this car dealership.

Overview

1. Introduction to ANOVA
2. One-Way ANOVA
   - Example
   - One-Way ANOVA Table
3. ANOVA Assumptions
   - Normality
   - Constant Variance
   - Remedial Measures
4. Two-Way ANOVA
5. Post Hoc Comparisons
6. Random Effects

Post-Hoc Comparisons

ANOVA compares all individual mean differences simulatneously. A significant result implies that at least one difference in means is statistically significant.

- However, ANOVA alone tells us nothing about which means differ significantly from one another.
- Post-Hoc tests are follow-up tests conducted to determine exactly which mean differences are significant, and which are not.

Post-Hoc Comparison

Back to the example presented for One-Way ANOVA:

- We know that cholesterol varies significantly with dose. That is, if μ_i is the mean cholesterol for dose i, μ_1, μ_2, and μ_3 are not all the same number.
- However, we do not know which of these means are different from the others.

Post-Hoc Comparison

In this case, a **multiple comparison** should be done. We would like to individually compare:

- μ_1 and μ_2
- μ_1 and μ_3
- μ_2 and μ_3

However, when doing multiple comparisons our proabability of falsely rejecting H_0 increases. If $\alpha = 0.05$, then *each comparison* will have a 5% chance of false rejection. Thus, a simultaneous statement about the three comparisons above will be incorrect with probability greater than 5%.

Post-Hoc Comparison

A good testing procedure should take into account this inflated error for multiple comparisons. The following are common methods of multiple comparison:

- Least Significant Difference (LSD)
- Tukey's Test
- Bonferroni Test
- Scheffé's Test

Each of these tests have different strengths and weaknesses.

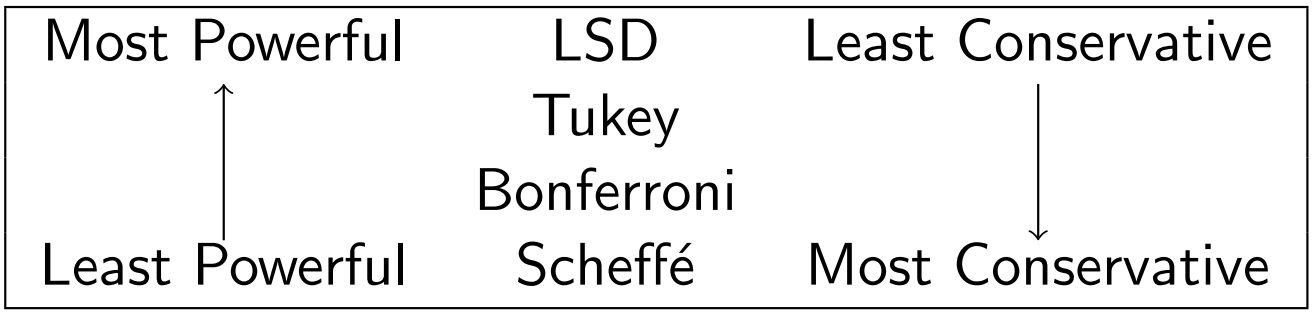
Post-Hoc Comparison

Each of these tests have different strengths and weaknesses.

- The **Bonferroni Test** is most useful when we are only interested in a small number of planned comparisons.
- **Tukey's Test** is used when we are only interested in performing all (or most) pairwise tests (i.e., tests of the form $H_0 : \mu_i = \mu_j,\ i \neq j$)
- The **Scheffé Test** is used when doing anything that might be considered *data snooping* (i.e., unplanned comparisons)

Post Hoc Comparison

There is a tradeoff between a test being powerful (having a large number of correct rejections) and conservative (having a small number of incorrect rejections):



Post-Hoc Test

We can run all the aforementioned post-hoc tests in R:

```
library(stats) #TukeyHSD
library(DescTools) #ScheffeTest

## Warning:  package 'DescTools' was built under R version
3.6.2

model <- aov(dose[,1] ~ dose[,2])
```

Post-Hoc Test

We can run all the aforementioned post-hoc tests in R:

```
pairwise.t.test(dose$dose, dose$group,
                p.adjust.method = "bonferroni")

##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  dose$dose and dose$group
##
##        group1 group2
## group2 1.000  -
## group3 0.051  0.179
##
## P value adjustment method: bonferroni
```

Post-Hoc Test

We can run all the aforementioned post-hoc tests in R:

```
TukeyHSD(aov(dose[,1] ~ dose[,2]))

##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = dose[, 1] ~ dose[, 2])
##
## $`dose[, 2]`
##                  diff       lwr       upr     p adj
## group2-group1  -12 -58.20875 34.208754 0.7720679
## group3-group1  -48 -94.20875 -1.791246 0.0416601
## group3-group2  -36 -82.20875 10.208754 0.1362300
```

Post-Hoc Test

We can run all the aforementioned post-hoc tests in R:

```
ScheffeTest(model)

##
##   Posthoc multiple comparisons of means: Scheffe Test
##     95% family-wise confidence level
##
## $`dose[, 2]`
##                diff    lwr.ci     upr.ci   pval
## group2-group1   -12 -60.28226 36.2822566 0.7903
## group3-group1   -48 -96.28226  0.2822566 0.0514 .
## group3-group2   -36 -84.28226 12.2822566 0.1580
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

Overview

1. Introduction to ANOVA
2. One-Way ANOVA
   - Example
   - One-Way ANOVA Table
3. ANOVA Assumptions
   - Normality
   - Constant Variance
   - Remedial Measures
4. Two-Way ANOVA
5. Post Hoc Comparisons
6. Random Effects

Random Effects Models

We have seen that Two-Way ANOVA can generalize the number of treatments in a study. In fact, we can conduct *n*-way ANOVA using the same principles (and increasingly difficult SS formulas!)

But *n*-way ANOVA still requires that our observations are independent.

- This may be an unreasonable assumption, especially if we are taking repeated measurements on a single subject. In this case, we would expect observations taken within each subject to be *dependent*

Random Effects Models

We can address this by incorporating a random quantity at the subject level. For instance, if Y_{ij} is the jth observation for the ith subject, then we could assume

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

Where $\epsilon_{ij} \overset{i.i.d}{\sim} \mathrm{N}\left(0, \sigma^2_\epsilon\right)$, $\tau_i \overset{i.i.d}{\sim} \mathrm{N}\left(0, \sigma^2_\tau\right)$, and τ_i's are independent of the ϵ_{ij}'s.

Random Effects Models

This setup assumes that $E(Y_{ij}) = 0$, $\text{Var}(Y_{ij}) = \sigma^2_\epsilon + \sigma^2_\tau$ and

- $\text{Cov}(Y_{ij}, Y_{ij'}) = \sigma^2_\tau$ for $j \neq j'$
- $\text{Cov}(Y_{ij}, Y_{i'j}) = 0$ for $i \neq i'$

Thus, the model assumes that there is correlation within subjects, but not across subjects

Random Effects Models

Random Effects models are typically fit in R using the `lmer()` or `lme4()` functions.

Credit

This workshop is based on a similar workshop delivered by Ruihan Lu in Fall 2020.

Services

Take Advantage of GradQuant's Services:

- GradQuant offers individual consultations through Skype. You are always welcome to make an appointment with us.
- Weekly Drop-in Hours on Wednesday (12pm - 2pm). Grad Students and Postdocs can meet with GradQuant consultants without an appointment, on a first-come, first-served basis.
- Weekly Hacky Hours on Monday (10am - 12pm). Hacky Hours are open to the whole campus, serving undergraduate and graduate students, faculty, and staff. No appointment is required.
- For detailed information on how to utilize our services, please visit: https://gradquant.ucr.edu

Questions?