****

**One Way ANOVA - Example**

A pharmaceutical company conducts an experiment to test the effect of a new cholesterol medication. The company selects 15 subjects randomly from a larger population. Each subject is randomly assigned to one of three treatment groups. Within each treament group, subjects receive a different dose of the new medication. In Group 1, subjects receive 0 mg/day; in Group 2, 50 mg/day; and in Group 3, 100 mg/day.

The treatment levels represent all the levels of interest to the experimenter, so this experiment used a [fixed-effects model](https://stattrek.com/statistics/dictionary?definition=fixed-effects-model) to select treatment levels for study.

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

|  |  |  |
| --- | --- | --- |
| **Dosage** | | |
| **Group 1, 0 mg** | **Group 2, 50 mg** | **Group 3, 100 mg** |
| 210 | 210 | 180 |
| 240 | 240 | 210 |
| 270 | 240 | 210 |
| 270 | 270 | 210 |
| 300 | 270 | 240 |

In conducting this experiment, the experimenter had two research questions:

* Does dosage level have a significant effect on cholesterol level?
* How strong is the effect of dosage level on cholesterol level?

**Solution**

Given Data

|  |  |  |
| --- | --- | --- |
| **Dosage** | | |
| **Group 1, 0 mg** | **Group 2, 50 mg** | **Group 3, 100 mg** |
| 210 | 210 | 180 |
| 240 | 240 | 210 |
| 270 | 240 | 210 |
| 270 | 270 | 210 |
| 300 | 270 | 240 |

**Step 1:**

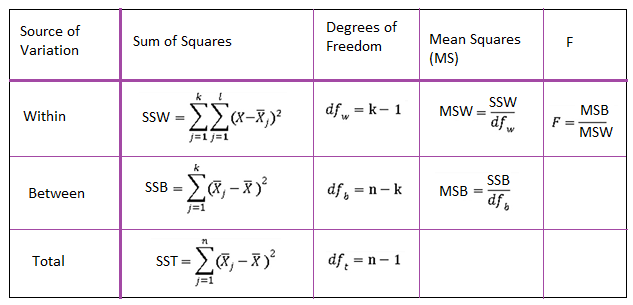
State Hypothesis

**Null Hypothesis**  : Ho : Mean\_Group1 = Mean\_Group2 = Mean\_Group3 – 0mg, 50mg,100mg has same effect on Cholesterol Level

**Alternate Hypothesis** : H1 : Mean\_Group1 ≠ Mean\_Group2 ≠ Mean\_Group3 - 0mg, 50mg,100mg does not same effect on Cholesterol Level

**Step 2:**

Summarize the One-Way ANOVA data



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Dosage** | | |  |  |  |
| **Group 1,** | **Group 2,** | **Group 3,** |  |  |  |
| **0 mg** | **50 mg** | **100 mg** | **(Group1 - Mean\_Group1)2** | **(Group2 - Mean\_Group2)2** | **(Group2- Mean\_Group2)2** |
| 210 | 210 | 180 | 2304 | 1296 | 900 |
| 240 | 240 | 210 | 324 | 36 | 0 |
| 270 | 240 | 210 | 144 | 36 | 0 |
| 270 | 270 | 210 | 144 | 576 | 0 |
| 300 | 270 | 240 | 1764 | 576 | 900 |
| Mean\_Group1 | Mean\_Group2 | Mean\_Group3 | sum= | sum= | sum= |
| 258 | 246 | 210 | 4680 | 2520 | 1800 |
| Grand Mean | 238 |  |  |  |  |
| Number of variables in Group1 | 5 | SS\_total | 15240 | df\_total | 14 |
| Number of variables in Group2 | 5 | SS\_between | 6240 | df\_between | 2 |
| Number of variables in Group3 | 5 | SS\_within | 9000 | df\_within | 12 |
| Total Number of Variables | 15 |  |  |  |  |
| Number of Groups | 3 | MSW | 750 | MSB | 3120 |
|  |  |  |  |  |  |
|  | F(Statistics) | 4.16 |  |  |  |

F critical value is = 3.89

**Step 3: Compare Statistics and Crtitical value**

Fctitical < Fstatistics

Step 4: Conclusion

Reject the Null Hypothesis

Accept Alternate Hypothesis

Mean\_Group1 ≠ Mean\_Group2 ≠ Mean\_Group3 - 0mg, 50mg,100mg does not same effect on Cholesterol Level is accepted

**R Code**

> g1<-c(210,240,270,270,300)

> g2<-c(210,240,240,270,270)

> g3<-c(180,210,210,210,240)

> grp1<-rep("0mg",5)

> grp2<-rep("50mg",5)

> grp3<-rep("100mg",5)

> df<-data.frame(drug\_level=c(grp1,grp2,grp3),cholestrol=c(g1,g2,g3))

> df

drug\_level cholestrol

1 0mg 210

2 0mg 240

3 0mg 270

4 0mg 270

5 0mg 300

6 50mg 210

7 50mg 240

8 50mg 240

9 50mg 270

10 50mg 270

11 100mg 180

12 100mg 210

13 100mg 210

14 100mg 210

15 100mg 240

> aov(df$cholestrol~df$drug\_level)

Call:

aov(formula = df$cholestrol ~ df$drug\_level)

Terms:

df$drug\_level Residuals

Sum of Squares 6240 9000

Deg. of Freedom 2 12

Residual standard error: 27.38613

Estimated effects may be unbalanced

> summary(aov(df$cholestrol~df$drug\_level))

Df Sum Sq Mean Sq F value Pr(>F)

df$drug\_level 2 6240 3120 4.16 0.0424 \*

Residuals 12 9000 750

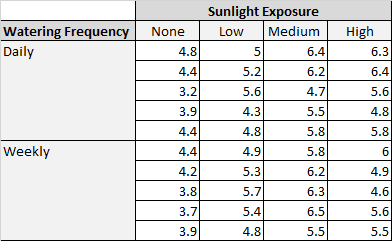
---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**Two Way ANOVA – Example**

Suppose a botanist wants to know if plant growth is influenced by sunlight exposure and watering frequency. She plants 40 seeds and lets them grow for one month under different conditions for sunlight exposure and watering frequency.

After one month, she records the height of each plant. The results are shown below:



In the table above, we see that there were five plants grown under each combination of conditions. Perform a Two-Way ANOVA analysis to determine which combination is better for plant growth.

**Solution**

**Step 1 : Identify the Groups**

Total Groups

Daily None, Daily Low, Daily Medium, Daily High

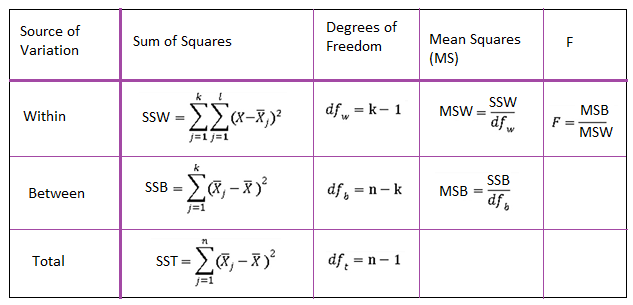
Weekly None, Weekly Low, Weekly Medium, Weekly High

Step 2: State Hypothesis

Null Hypothesis : Ho : No significant difference in mean of the groups

Alternate Hypothesis : H1 : There is significant difference in mean of different groups

**Step 2: Summarize the Two-Way ANOVA table**



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Sun Light Exposure | | | |
| Watering Frequency | None | Low | Medium | High |
| Daily | 4.8 | 5 | 6.4 | 6.3 |
| 4.4 | 5.2 | 6.2 | 6.4 |
| 3.2 | 5.6 | 4.7 | 5.6 |
| 3.9 | 4.3 | 5.5 | 4.8 |
| 4.4 | 4.8 | 5.8 | 5.8 |
| Weekly | 4.4 | 4.9 | 5.8 | 6 |
| 4.2 | 5.3 | 6.2 | 4.9 |
| 3.8 | 5.7 | 6.3 | 4.6 |
| 3.7 | 5.4 | 6.5 | 5.6 |
| 3.9 | 4.8 | 5.5 | 5.5 |
| Mean | 4.07 | 5.1 | 5.89 | 5.55 |
| Grand Growth Mean | 5.1525 |  |  |  |
| Daily Watering Growth Mean | 5.155 |  |  |  |
| Weekly Watering Growth Mean | 5.15 |  |  |  |
| Sum of Squares of Watering Frequency | 20(5.155-5.1525)^2 + 20(5.15-5.1525)^2 =0.00025 | | |  |
| Sum of Squares of Sun Light Exposure | 10(4.07-5.1525)^2 + 10(5.1-5.1525)^2 + 10(5.89-5.1525)^2 + 10(5.55-5.1525)^2 ==18.76475=18.76475 | | |  |
| Total Sum of Squares | 28.45975 |  |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sum Squares Daily None | 1.512 |  | df Water | 2-1 =1 | j=2 | |  | |
| Sum Squares Daily Low | 0.928 |  | df Sun Light | 4-1=3 | k=4 | |  | |
| Sum Squares Daily Medium | 1.788 |  | df interaction | 1\*3 =3 | j-1\*k-1 | |  | |
| Sum Squares Daily High | 1.648 |  | df within | 40 – (2\*4) = 32 | n – (j\*k) | |  | |
| Sum Squares Weekly None | 0.34 |  |  |  |  | |  | |
| Sum Squares Weekly Low | 0.548 |  | F Watering | MSWatering/MSWithin |  | | 0.000288 | |
| Sum Squares Weekly Medium | 0.652 |  | F SunLight | MSSunLight/MSWithin |  | | 2.160842 | |
| Sum Squares Weekly High | 1.268 |  | F Interaction | MSInteraction/MSWithin |  | | 0.116392 | |
| Sum Squares within | 8.684 |  |  |  |  | |  | |
|  |  |  |  |  |  | |  | |
|  |  |  |  |  |  | |  | |
| SS Interaction = SS Total – SS Factor 1 – SS Factor 2 – SS Within= 28.45975 – .00025 – 18.76475 – 8.684 = 1.01075 | | | | | |  | |
|  | |

**Step 3: Conclusion**F Watering, F SunLight, F Interaction statistics values are significant than F Critical values

It is concluded that Sun Light and Watering has effect on plant growth

**R Code**

> df<-read.csv("C:/Users/User/Documents/plant\_growth.csv")

> df

Water Sun Growth

1 D N 4.8

2 D N 4.4

3 D N 3.2

4 D N 3.9

5 D N 4.4

6 D L 5.0

7 D L 5.2

8 D L 5.6

9 D L 4.3

10 D L 4.8

11 D M 6.4

12 D M 6.2

13 D M 4.7

14 D M 5.5

15 D M 5.8

16 D H 6.3

17 D H 6.4

18 D H 5.6

19 D H 4.8

20 D H 5.8

21 W N 4.4

22 W N 4.2

23 W N 3.8

24 W N 3.7

25 W N 3.9

26 W L 4.9

27 W L 5.3

28 W L 5.7

29 W L 5.4

30 W L 4.8

31 W M 5.8

32 W M 6.2

33 W M 6.3

34 W M 6.5

35 W M 5.5

36 W H 6.0

37 W H 4.9

38 W H 4.6

39 W H 5.6

40 W H 5.5

> aov(Growth~Water+Sun,data=df)

Call:

aov(formula = Growth ~ Water + Sun, data = df)

Terms:

Water Sun Residuals

Sum of Squares 0.00025 18.76475 9.69475

Deg. of Freedom 1 3 35

Residual standard error: 0.5263011

Estimated effects may be unbalanced

> summary(aov(Growth~Water+Sun,data = df))

Df Sum Sq Mean Sq F value Pr(>F)

Water 1 0.000 0.000 0.001 0.976

Sun 3 18.765 6.255 22.582 2.59e-08 \*\*\*

Residuals 35 9.695 0.277

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1