Lecture 19

Part 4 Analysis of Variance

Two-way ANOVA

Two-way ANOVA

- In a two-way ANOVA, we have:
  - One continuous response variable Y.
  - Two categorical variables (factors), each with a potentially different number of categories (levels).
  - Each factor must still have at least two levels.
- For a two-way ANOVA, a particular combination of levels from the two factors is called a **treatment** and represents a population.

Treatment

- Suppose our two factors, denoted A and B, have three levels (A_1, A_2, A_3) and two levels (B_1, B_2), respectively.
- There are $3 \times 2 = 6$ treatments, as shown below:

Treatment	Levels	
1	A_1	B_1
2	A_1	B_2
3	A_2	B_1
4	A_2	B_2
5	A_3	B_1
6	A_3	B_2

Terminology

- A **complete factorial experiment** is one where sample data is collected for every possible combination of levels of the two factors. That is, we have data for all treatments.
- A complete factorial experiment is **balanced** if the number of observations collected for each treatment (also called **replicates**) is the same.
- In the example on the previous slide, if we collected, e.g., five replicates for each of the six treatments, we would have a balanced two-way ANOVA.

Assumptions

- We must also make some assumptions when performing a two-way ANOVA:

1. The levels of both factors are fixed beforehand.
2. The response variable is normally distributed with constant variance in each treatment.
3. Samples are independent.

Two-way ANOVA

- A two-way ANOVA allows us to ask and answer more interesting questions than a one-way ANOVA.
- Letting A and B again denote our two factors, we can use a two-way ANOVA to answer:
  - Does the mean response change for different levels of factor A?
  - Does the mean response change for different levels of factor B?
  - Do the factors A and B interact?

Interaction

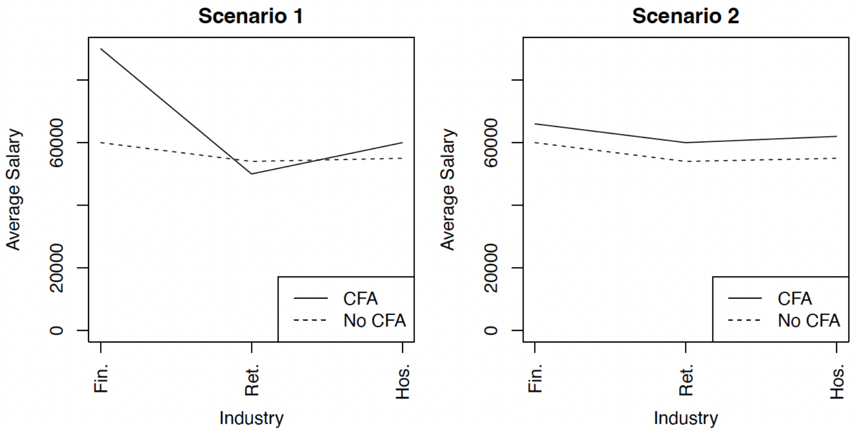
- Suppose we want to analyze new graduate salaries, based on qualification and industry.
- Some sample mean starting salaries of graduates with and without the Chartered Financial Analyst (CFA) credential, in the finance, retail and hospitality industries, are shown over.

Interaction

Industry	Scenario 1		Scenario 2	
	CFA	No CFA	CFA	No CFA
Finance	90K	60K	66K	60K
Retail	50K	52K	60K	52K
Hospitality	60K	65K	72K	65K

- Potential interactions are easiest to identify by graphing the sample means.

Interaction



Interaction

- In scenario 1, the impact of having a CFA is most pronounced in the finance industry.
- In scenario 2, the impact of having a CFA is the same regardless of industry.
- We say that two factors **interact** when the effect of one factor on the response variable is altered by the level of the other factor.
- So there is an interaction between qualification and industry in scenario 1 but not in scenario 2.

Interaction Hypotheses

- With a two-way ANOVA we can test whether an interaction exists between the factors:

  H_0 : There is *no interaction* between the factors.

  H_1 : There is *an interaction* between the factors.

- The interaction hypotheses should be tested first, before testing the *main effects hypotheses*.

Main Effects Hypotheses

- With a two-way ANOVA we can test the importance of each individual factor:

  H_0 : The population means at different levels of the factor are all equal.

  H_1 : At least two of the population means differ.
- If an interaction exists, the results of the main effects hypotheses should be interpreted carefully.

Sums of Squares

- Suppose our two factors, denoted A and B, have a and b levels, respectively, and we sample r replicates in each treatment (i.e., $n = a \times b \times r$).
- We can calculate a sum of squares for factor A (SS_A), for factor B (SS_B), for the interaction (SS_{AB}), for the error (SSE) and also for the total ($SS(Total)$).

Sums of Squares

- The sums of squares satisfy the following identity:

$$SS(Total) = SS_A + SS_B + SS_{AB} + SSE$$

- The $SS(Total)$ again measures the total variation that exists in the data.
- The SS_A, SS_B and SS_{AB} measure how much variation can be explained by each particular source.
- The SSE measures the left-over, unexplained variation.

Test Statistics

- To test the interaction and main effects hypotheses, we need to see how big the SS_{AB}, SS_A and SS_B are, compared to the SSE.
- Again convert the sums of squares to mean squares, so that we can calculate F-statistics.
- The appropriate degrees of freedom, the mean squares and the F-statistics are listed in the ANOVA table shown next.

ANOVA Table

Source	Sum of squares	Degrees of freedom	Mean squares	F-statistic
Factor A	SS_A	$a-1$	$MS_A = \frac{SS_A}{a-1}$	$F_A = \frac{MS_A}{MSE}$
Factor B	SS_B	$b-1$	$MS_B = \frac{SS_B}{b-1}$	$F_B = \frac{MS_B}{MSE}$
Interaction	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$F_{AB} = \frac{MS_{AB}}{MSE}$
Error	SSE	$n-ab$	$MSE = \frac{SSE}{n-ab}$	
Total	$SS(Total)$	$n-1$		

ANOVA Table - Summary

- Degrees of freedom:
    - For a factor it is equal to the number of levels minus one.
    - For the interaction it is equal to the product of the degrees of freedom of the two factors.
    - All the degrees of freedom (excluding the total) will sum to $n - 1$.
- F-statistic:
    - Mean squares are calculated by dividing the sum of squares by the degrees of freedom.
    - All F-statistics use the MSE in the denominator.

Decision Rule

- We need to compare each F-statistic to an appropriate F-distribution.
- All hypotheses (interaction and main effects) are *one-tailed*, and we reject H_0 if the F-statistic is too large.
- At a significance level of α, we reject H_0 if $F > F_{\alpha,df,n-ab}$, where df is the source degrees of freedom and $F_{\alpha,df,n-ab}$ is the critical value that cuts off $100\alpha\%$ in the upper tail of an F-distribution with df numerator degrees of freedom and $n - ab$ denominator degrees of freedom.

Airlines Rating Example

- Recall the previous example where twenty passengers, from each of three airlines, rated their experience on a 0 to 100 scale.
- Suppose that for each airline, half of the twenty passengers traveled in business class, while the remaining half traveled in economy class.
- We want to know:
  - Does perceived quality vary between airlines?
  - Does perceived quality vary between traveling class?
  - Does traveling class alter the differences between airlines in terms of perceived quality?

ANOVA Table

Source	Sum of squares	Deg. of freedom	Mean squares	F-statistic	p-value
Airline	3446.80	2	1723.40	14.60	0.0000
Class	6060.15	1	6060.15	51.34	0.0000
Interaction	130.00	2	65.00	0.55	0.5798
Error	6374.30	54	118.04		
Total	16011.25	59			

Interaction

- Hypotheses:

  H_0 : Airline and class do not interact.
  H_1 : Airline and class interact.

- From the output, we see that the F-statistic is 0.55 with a p-value of 0.5798.
- Since the $0.5798 > 0.05$, we fail to reject H_0 and we conclude that there is no interaction between airline and traveling class.

Main Effects

- Hypotheses:

$$H_0 : \text{Pop. mean ratings are equal between airlines.}$$
$$H_1 : \text{Pop. mean ratings differ between airlines.}$$

- From the output, we see that the F-statistic is 14.60 with a p-value of 0.0000.
- Since the $0.0000 < 0.05$, we reject H_0 and we conclude that perceived quality varies between airlines.

Main Effects

- Hypotheses:

$$H_0 : \text{Pop. mean ratings are equal between class.}$$
$$H_1 : \text{Pop. mean ratings differ between class.}$$

- From the output, we see that the F-statistic is 51.34 with a p-value of 0.0000.
- Since the $0.0000 < 0.05$, we reject H_0 and we conclude that perceived quality varies between travelling class.

Airline Ratings ANOVA Tables

	Source	Sum of squares	Deg. of freedom	Mean squares	F-statistic
One-way ANOVA	Airline	3446.8	2	1723.4	7.8184
	Error	12564.45	57	220.4289	
	Total	16011.25	59		
Two-way ANOVA	Airline	3446.80	2	1723.40	14.60
	Class	6060.15	1	6060.15	51.34
	Interaction	130.00	2	65.00	0.55
	Error	6374.30	54	118.04	
	Total	16011.25	59		

One-way and Two-way ANOVA

- Suppose we perform both a one-way ANOVA and a two-way ANOVA, using a *balanced design*, on the *same* data with a factor in common.
  - The $SS(\text{Total})$ will be the same for both.
  - The SST in the one-way ANOVA will be the same as the sum of squares for the corresponding factor in the two-way ANOVA.
  - The two-way ANOVA accounts for more variation than the one-way ANOVA.
  - This means that for the two-way ANOVA, there is *less* unexplained variation so the SSE will be smaller than for the one-way ANOVA.