Part-A: Two-way ANOVA

LEARNING OBJECTIVES

At the end of this section, you should be able to do the following:

- Design and conduct a two-way ANOVA experiment
- Interpret and explain the main effect and interaction effect of the factors
- Illustrate the cell means plots for interaction effect

WHY TWO WAY ANOVA?

- In the previous topic, we used one-way analysis of variance to test for differences between three or more population means.
- The one-way examples include comparisons by gender, ethnicity, political party, etc.)
- In this topic we extend one-way ANOVA and introduce two factors (two factors or two levels).
- It answers questions such as:
  - Are the means of the population corresponding to factors different?
  - Are the means corresponding to one factor different to another?
  - Do the factors interact?

UNDERSTANDING TWO WAY ANOVA

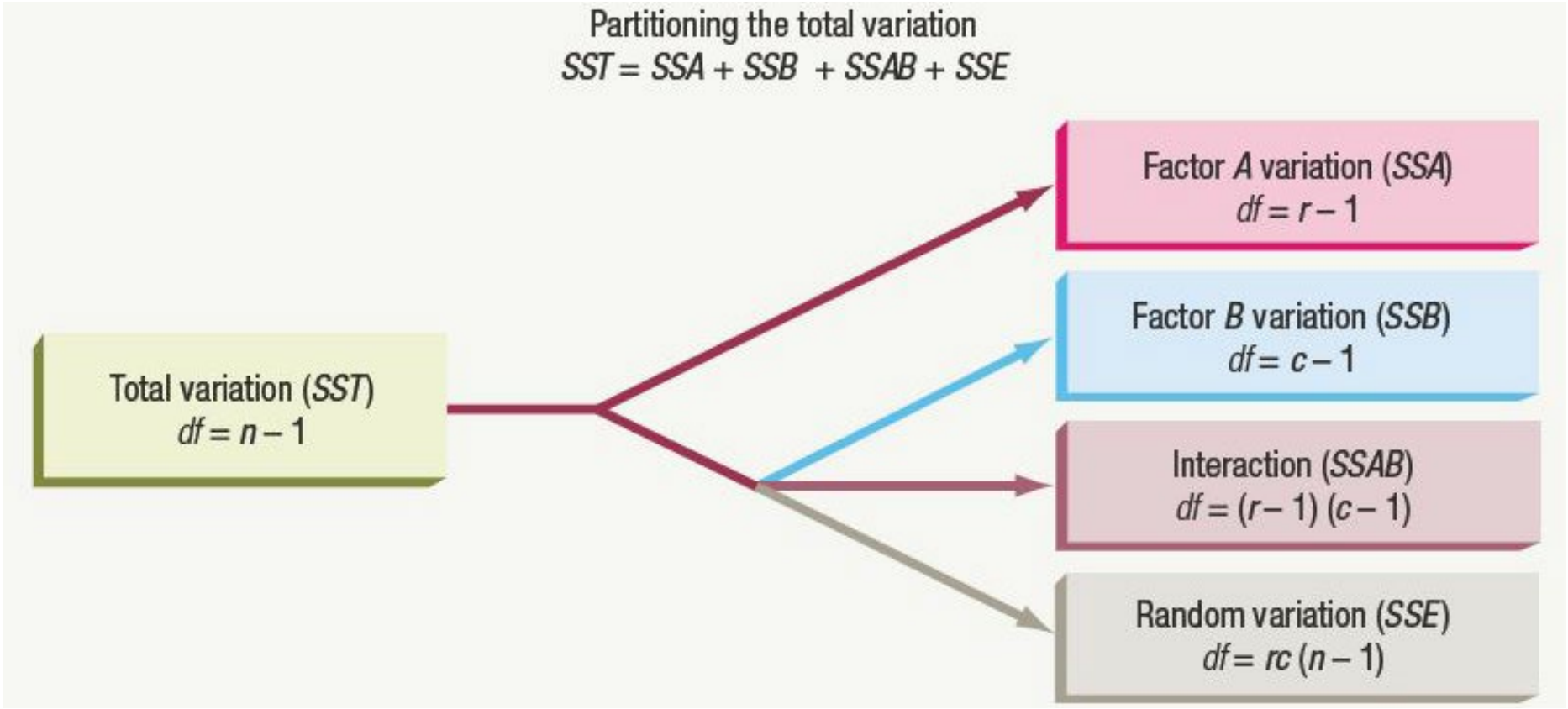
- Vocabulary
  - Response = Dependent Variable (DV) (numerical)
  - Factor = Independent Variable (IV) (categorial)
  - Rows = Number of Levels of Factor A
  - Columns = Number of Levels of Factor B

We organise data in a table format

- Understanding calculations
  - The calculation for Sums of Squares for a two-way ANOVA by hand is time consuming.
  - In practice we almost of exclusively use a statistical package.
  - But you need to have a conceptual understanding of the calculation

UNDERSTANDING TWO WAY ANOVA

- Partitioningthe total variation in a two-way ANOVA



UNDERSTANDING TWO WAY ANOVA

- Sum of Squares for Factor A
  - SSA Measures variation in the response due to the fact that different levels of factor A were used.
- Sum of Squares for Factor B
  - SSB Measures variation in the response due to the fact that different levels of factor B were used.
- Interaction Sum of Squares
  - SSAB Measures the variation in the response due to the interaction between factors A and B. If the interaction plot is perfectly parallel this will be 0.
- Error or Residual Sum of Squares
  - SSE Measures the variation in the response within

  the a x b factor combinations.

UNDERSTANDING TWO WAY ANOVA

- There are three distinct tests to perform in a-Two-way ANOVA
  - To test the hypothesis of no difference due to factor A:
    - H0: $\mu 1..= \mu 2.. = \mu 3..= \ldots = \mu r..$
    - H1: Not all $\mu i..$ are equal
    - Reject H_0 if $F_{STAT} > F_u$
  - To test the hypothesis of no difference due to factor B:
    - H0: $\mu .1. = \mu .2. = \mu .3.= \ldots = \mu .c.$
    - H1: Not all $\mu .j.$ are equal
    - Reject H_0 if $F_{STAT} > F_u$
  - To test the hypothesis of no interaction of factors A and B:
    - H0: the interaction of A and B is equal to zero
    - H1: interaction of A and B is not zero
    - Reject H_0 if $F_{STAT} > F_u$

UNDERSTANDING TWO WAY ANOVA

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F	p value
FactorA	SSA	r – 1	MSA = SSA/(r – 1)	$\frac{MSA}{MSE}$	Tail area
Factor B	SSB	c – 1	MSB = SSB/(c – 1)	$\frac{MSB}{MSE}$	Tail area
AB(Interaction)	SSAB	(r-1)(c-1)	MSAB = SSAB/ (r – 1)(c – 1)	$\frac{MSAB}{MSE}$	Tail area
Error	SSE	rc(n' – 1)	MSE = SSE/ rc(n' – 1)		
Total	SST	n – 1			

This is our initial focus which is the p-value for Question 1: Is there an interaction effect?

UNDERSTANDING TWO WAY ANOVA

- Observational data: donation ($)
- We have to reorganise data to reflect the two factors

Homeless Person	*Donor& Friends*		
	0	*1*	*2*
Female	5	10	10
	3	8	10
	2	9.5	9
	3	10	8
	2	2.5	7
Male	1.5	3	5
	2	4	4
	2.5	5	3
	2	4	4
	2	4	4

UNDERSTANDING TWO WAY ANOVA

- We are interested in the generosity towards homeless people (DV = $ donated) in the following situations.
  - When the homeless person is male or female
  - When the donor is alone or with 1 or 2 other friends
- IV_1 : The homeless person is either male or female
- IV_2 : The donor is alone, or with 1 OR 2 friends

- How many levels do we have? We have five levels (two for gender, and three for donors)
- How many treatment combinations? 2 X 3 = 6

UNDERSTANDING TWO WAY ANOVA

- Main Effects are:
  - The effect of One IV on the DV (donation: X=$) averaged across the levels of the other IV.
- Main effect of gender:
  - Is there a difference in donation if the homeless person is male versus female, averaging over the number of friends present?
  - Ignoring the number of friends, does the donations differ for male versus female homeless persons?
- Main effects if number of friends presence:
  - Is there a difference in donations with 1 or 2 friends presents, averaging over the homeless person's gender?
  - Ignoring the gender of the homeless person, does the donations differ based on the number of friends present?

UNDERSTANDING TWO WAY ANOVA

	Donor + 2 Friends (n=10)	Donor + 1 Friends (n=10)	Donor + 0 Friends (n=10)
Female (n=15)	$\overline{X} = 9$	$\overline{X} = 8$	$\overline{X} = 3$
Male (n=15)	$\overline{X} = 4$	$\overline{X} = 4$	$\overline{X} = 2$

- The Two-Way ANOVA offers three different statistical tests:
  - Main effect of IV_1: Gender of Homeless person
  - Main effect of IV_2: Number of Friends present
  - Interaction between the two IVs (Gender of Homeless person and Number of Friends Present)

UNDERSTANDING TWO WAY ANOVA

Hypothesis Tests:

- Gender Main Effect:
  - H_0: Mean donation received is no different between male and female homeless person
  - H_1: Mean donation received is different between male and female homeless person
- Friends Presence Main Effect:
  - H_0: Mean donated amount does not differ with the number of friends present
  - H_1: Mean donated amount does differ with the number of friends present
- Interaction Effects:
  - H_0: there is no interaction between factors (Homeless person's gender and the number of Friends present)
  - H_1: there is an interaction between factors (Homeless person's gender and the number of Friends present)

UNDERSTANDING TWO WAY ANOVA

- Step 1: Computing marginal means to understand Main Effects

	Donor + 2 Friends (n=10)	Donor + 1 Friends (n=10)	Donor + 0 Friends (n=10)	Margin Means
Female (n=15)	$\bar{X} = 9$	$\bar{X} = 8$	$\bar{X} = 3$	$\bar{X} = 6.67$
Male (n=15)	$\bar{X} = 4$	$\bar{X} = 4$	$\bar{X} = 2$	$\bar{X} = 3.33$
Margin Means	$\bar{X} = 6.5$	$\bar{X} = 6$	$\bar{X} = 2.5$	

Response for female homeless person, average over # in party (9+8+3)/3 = 6.67

Response: party of 3, averaging over gender (9 + 4) / 2 = 6.5

Response: party of 1, averaging over gender (3 + 2) / 2 = 2.5

UNDERSTANDING TWO WAY ANOVA

- Step 2: Calculations: Main Effects Sums of Squares
  - Calculations for main effects SS in a Two-Way ANOVA are very similar to the calculations we used in one-way ANOVA
  - Conceptually, to calculate the SS for a main effect, one is comparing each marginal to the overall (Grand) mean

	Donor + 2 Friends (n=10)	Donor + 1 Friends (n=10)	Donor + 0 Friends (n=10)	Margin Means
Female (n=15)	$\bar{X} = 9$	$\bar{X} = 8$	$\bar{X} = 3$	$\bar{X} = 6.67$
Male (n=15)	$\bar{X} = 4$	$\bar{X} = 4$	$\bar{X} = 2$	$\bar{X} = 3.33$
Margin Means	$\bar{X} = 6.5$	$\bar{X} = 6$	$\bar{X} = 2.5$	Overall: 5

UNDERSTANDING TWO WAY ANOVA

- Main effect of gender question:
  - Do the marginal means of 6.67 (female homeless persons) and 3.33 (male homeless persons) differ?
- Main effect of Number of Friends Presence question:
  - Do the marginal means of 6.5 (2 friends), 6 (1 friend), and 2.5 (No friends) differ?

UNDERSTANDING TWO WAY ANOVA

- Calculations: Main Effects: gender
  1. Take the squared difference
  - The mean from female homeless persons & the overall mean (6.67-5)
  - The mean from male homeless persons & the overall mean (3.33-5)
  2. $SS_{Gender} = 15(6.67-5)^2 + 15(3.33-5)^2$

	Donor + 2 Friends (n=10)	Donor + 1 Friends (n=10)	Donor + 0 Friends (n=10)	Margin Means
Female (n=15)	$\bar{X} = 9$	$\bar{X} = 8$	$\bar{X} = 3$	$\bar{X} = 6.67$
Male (n=15)	$\bar{X} = 4$	$\bar{X} = 4$	$\bar{X} = 2$	$\bar{X} = 3.33$
Margin Means	$\bar{X} = 6.5$	$\bar{X} = 6$	$\bar{X} = 2.5$	Overall: 5

UNDERSTANDING TWO WAY ANOVA

- Calculations: Main Effects: number of Friends
  1. Take the squared difference of b/n
  – The mean from 2 friends & the overall mean (6.5-5)
  – The mean from 1 friends & the overall mean (6-5)
  – The mean from 0 friends & the overall mean (2.5-5)
  2. $SS_{Friends} = 10(6.5-5)^2 + 10(6-5)^2 + 10(2.5-5)^2$

	Donor + 2 Friends (n=10)	Donor + 1 Friends (n=10)	Donor + 0 Friends (n=10)	Margin Means
Female (n=15)	$\bar{X} = 9$	$\bar{X} = 8$	$\bar{X} = 3$	$\bar{X} = 6.67$
Male (n=15)	$\bar{X} = 4$	$\bar{X} = 4$	$\bar{X} = 2$	$\bar{X} = 3.33$
Margin Means	$\bar{X} = 6.5$	$\bar{X} = 6$	$\bar{X} = 2.5$	Overall: 5

UNDERSTANDING TWO WAY ANOVA

- Calculations: Sums of Squares (SS) cells
  - Sum the squared deviations between each cell mean and the overall grand mean
  - Each deviation is weighted by the number of observations in that cell

$$SS_{cells} = 5\,(9-5)^2 + 5\,(8-5)^2 + 5\,(3-5)^2 + 5\,(4-5)^2 + 5\,(4-5)^2 + 5\,(2-5)^2$$

	Donor + 2 Friends (n=10)	Donor + 1 Friends (n=10)	Donor + 0 Friends (n=10)	Margin Means
Female (n=15)	$\bar{X} = 9$	$\bar{X} = 8$	$\bar{X} = 3$	$\bar{X} = 6.67$
Male (n=15)	$\bar{X} = 4$	$\bar{X} = 4$	$\bar{X} = 2$	$\bar{X} = 3.33$
Margin Means	$\bar{X} = 6.5$	$\bar{X} = 6$	$\bar{X} = 2.5$	Overall: 5

UNDERSTANDING TWO WAY ANOVA

- Calculations: Total Sums of Squares

Homeless Person	*Donor& Friends*		
	0	*1*	*2*
Female	5	10	10
	3	8	10
	2	9.5	9
	3	10	8
	2	2.5	7
Male	1.5	3	5
	2	4	4
	2.5	5	3
	2	4	4
	2	4	4

$(5-5)^2$	$(10-5)^2$	$(10-5)^2$
$(3-5)^2$	$(8-5)^2$	$(10-5)^2$
$(2-5)^2$	$(9.5-5)2$	$(9-5)^2$
$(3-5)^2$	$(10-5)^2$	$(8-5)^2$
$(2-5)^2$	$(2.5-5)^2$	$(7-5)^2$
$(1.5-5)^2$	$(3-5)^2$	$(5-5)^2$
$(2-5)^2$	$(4-5)^2$	$(4-5)^2$
$(2.5-5)^2$	$(5-5)^2$	$(3-5)^2$
$(2-5)^2$	$(4-5)^2$	$(4-5)^2$
$(2-5)^2$	$(4-5)^2$	$(4-5)^2$

UNDERSTANDING TWO WAY ANOVA

- Calculations: Sums of Squares Interaction
  - We calculated SS_{Gender}, $SS_{Friends}$ and SS_{T} directly
  - $SS_{Interaction} = SS_{cells} - SS_{Gender} - SS_{Friends}$

- Calculations: Sums of Squares Errors
  - The SS_{error} is "What's left over"
  - Of the SS_{total}, we know what is due to gender (SS_{Gender}), what is due to ($SS_{Friends}$) and what is due to the interaction ($SS_{Interaction}$). Thus
  - $SS_{error} = SS_{total} - (SS_{Gender} + SS_{Friends} + SS_{Interaction})$

UNDERSTANDING TWO WAY ANOVA

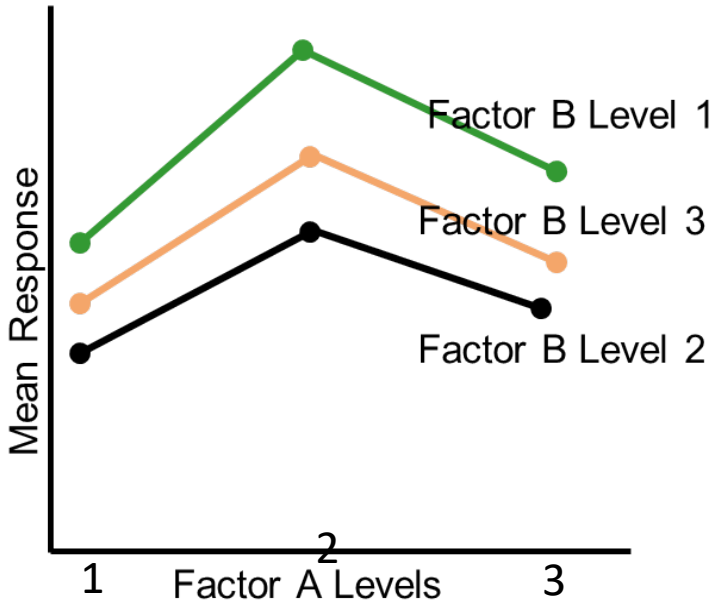
Anova: Two-Factor With Replication				
SUMMARY	Two	One	None	Total
Female				
Count	5	5	5	15
Sum	44	40	15	99
Average	8.8	8	3	6.6
Variance	1.7	10.13	1.5	10.86
Males				
Count	5	5	5	15
Sum	20	20	10	50
Average	4	4	2	3.333
Variance	0.5	0.5	0.125	1.274
Total				
Count	10	10	10	
Sum	64	60	25	
Average	6.4	6	2.5	
Variance	7.378	9.167	1	

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	80.03	1	80.03	33.23	0.000	4.26
Columns	92.07	2	46.03	19.11	0.000	3.403
Interaction	20.07	2	10.03	4.166	0.028	3.403
Within	57.8	24	2.408			
Total	250	29				

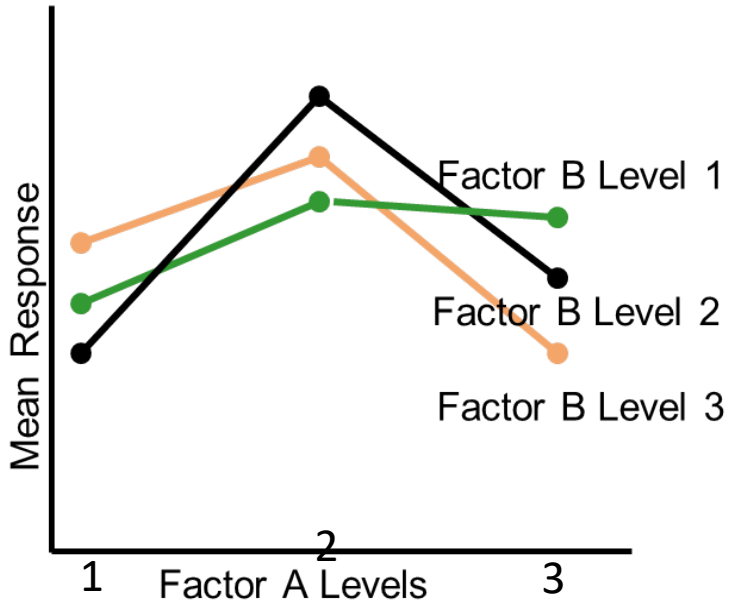
UNDERSTANDING TWO WAY ANOVA

- Cellmean plot to understand interaction better

No interaction: line segments are parallel

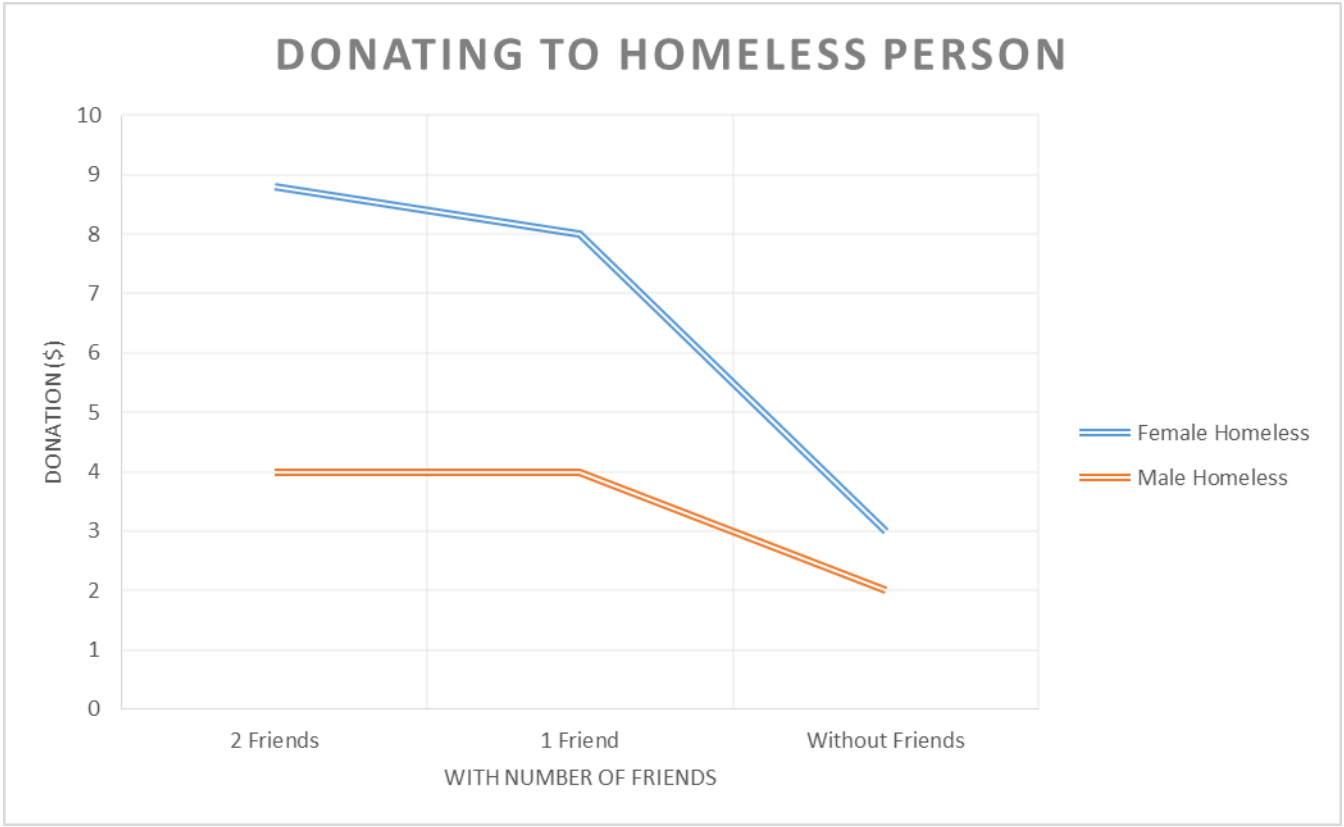


Interaction is present: some line segments not parallel



UNDERSTANDING TWO WAY ANOVA

- Cellmean plot to understand interaction better



UNDERSTANDING TWO WAY ANOVA

Interpretation of Two-way ANOVA

- The interaction effect:
  - p-value is 0.028 which is less than the $\alpha = 0.05$. We reject the H_0.
  - There is an interaction between the number of friends present and the gender of the homeless person.
  - Simple effects tests showed that male and female homeless persons receive almost equal donations when no friends of the donor were present.
  - However, when 1 or 2 friends of the donor were present, female homeless persons received larger donations than male homeless persons.

UNDERSTANDING TWO WAY ANOVA

Interpretation of Two-way ANOVA

- Gender Main Effect:
  - Do the marginal means 6.6 (female) and 3.3 (male) differ?
  - p-value is 0.00 which is less than the $\alpha = 0.05$. We reject the H_0.
  - There is sufficient evidence to conclude that female homeless persons receive larger donations than male homeless persons, on average.

Superseded because of the interaction effect: It does not hold true across the number of friends – when with friends donations higher than without friends for female homeless person.

UNDERSTANDING TWO WAY ANOVA

Interpretation of Two-way ANOVA

- Friends Presence Main Effect:
  - Do the marginal means 2.5 (Alone), 6(1 Friend), and 6.4 (2 Friend2) differ?
  - p-value is 0.00 which is less than the $\alpha = 0.05$. We reject the H_0.
  - There is sufficient evidence to conclude that when the donor with one or more friends give larger donations to homeless persons, on average.

Superseded because of the interaction effect

QUESTIONS?