MATH 2444: Statistical Data Analysis

Assignment # 5 (Due 07/11/2025 @ 2pm)
LATE SUBMISSIONS WILL NOT BE ACCEPTED

Name: Jordan Serko

Chapter 13: Two-Way Analysis of Variance

Multiple Choice & True/False Questions

1. Which of the following statements about a two-way analysis of variance model is/are TRUE?

A) The population of interest is classified according to two categorical variables, or factors.
B) An experiment involving the simultaneous use of two factors offers advantages over two one-way experiments with respect to such matters as efficiency and reduction of residual variation.
C) The two-way type of experiment requires twice as many experimental observations as would be required in two one-way experiments of the same factors.
D) An experiment involving the simultaneous study of two factors allows for the investigation of interactions between the factors.
E) Only A, B, and D are true.

2. Two-way ANOVA is used to study the effect of two different categorical factors on one quantitative response variable.
TRUE **FALSE**

3. Which of the following statements about the two-way ANOVA is/are FALSE?

A) A two-way ANOVA is used when there are two outcome variables.
B) In a 3 × 3 ANOVA each level of Factor A appears with only two levels of Factor B.
C) The two-way ANOVA can only be used when the samples sizes are the same in all cells.
D) The presence of interactions always tells us that the main effects are uninformative.
E) All of the above are false.

4. A study compares three levels of Factor A and two levels of Factor B, with five observations in each cell. What are the degrees of freedom for the *F* statistic that is used to test for interaction?

A) 2 and 24
B) 3 and 30
C) 5 and 6
D) 6 and 24

5. A study compares three levels of Factor A and four levels of Factor B, with seven observations in each cell. What are the degrees of freedom for the *F* statistic that is used to test for the main effect of Factor B?

A) 3 and 36

B) 3 and 72
C) 4 and 72
D) 4 and 84

6. A study compares two levels of Factor A and four levels of Factor B, with four observations in each cell. How large does the *F* statistic need to be for rejecting the null hypothesis that no interaction is present between the two factors at the 5% significance level? **Hint:** Use the **qf()** R function to find this value.

A) 2.33
B) 2.78
C) 2.92
D) 3.01

7. A study compares three levels of Factor A and four levels of Factor B, with three observations in each group. An *F* statistic of 5.113 is reported for the main effect of Factor A.

What can we say about the *P*-value for this *F* test? **Hint:** Use the 1 - **pf(** TS, df_1, df_2*)* R function to find this value.

A) *P*-value < 0.01
B) 0.01 <*P*-value < 0.025
C) 0.025 <*P*-value < 0.05
D) *P*-value > 0.05

8. In a study comparing four levels of Factor A and two levels of Factor B, with six observations in each cell, the SSE = 3456 and the SSA = 364. What is the value of the *F* statistic that is used to test for the main effect of Factor A?

A) 1.05
B) 1.26
C) 1.40
D) 4.21

9. In a two-way ANOVA, the interaction effect tells us whether the effect of one factor depends on the level of the other factor.
TRUE
FALSE

10. In a two-way ANOVA, the Mean Square Error (MSE) represents the variability of observations within treatment combinations and is used as the denominator in the F-tests for main and interaction effects.

TRUE

FALSE

11. If the interaction between two factors is not significant, we usually interpret only the main effects of the factors.
TRUE **FALSE**

12. In a two-way ANOVA, each factor has its own degrees of freedom based on the number of levels it has.
FALSE

TRUE

13. A research project studied the physical properties of wood materials constructed by bonding together small flakes of wood. The two factors considered were the size of the flakes and species of tree. The sizes of the flakes were S1: 0.15 inches by 2 inches and S2: 0.25 inches by 2 inches, and the species of tree used were aspen, birch, and maple. For each combination of flake size and tree species, three samples of wood material were constructed.

For each sample, the physical property measured was the tension modulus of elasticity in the direction perpendicular to the alignment of the flakes, in pounds per square inch (psi). Which statement best describes the type of experiment being run?

A) This is a two-factor experiment. Each factor is at three levels.
B) This is a six-level experiment. Each level has three observations.
C) This is a two-factor experiment. One factor is at two levels, and the other factor is at three levels.
D) None of the above

14. A research project studied the physical properties of wood materials constructed by bonding together small flakes of wood. The two factors considered were the size of the flakes and species of tree. The sizes of the flakes were S_1: 0.15 inches by 2 inches and S_2: 0.25 inches by 2 inches, and the species of tree used were aspen, birch, and maple. For each combination of flake size and tree species, three samples of wood material were constructed.

For each sample, the physical property measured was the tension modulus of elasticity in the direction perpendicular to the alignment of the flakes, in pounds per square inch (psi). The table below gives the means for the different size-species groups.

		Species		
		Aspen	**Birch**	**Maple**
Size	**S1**	387.33	292.67	323.33
	S2	335.67	455.33	293.67

What are the marginal means for species?

A) 387.33, 292.67, and 323.33
B) 335.67, 455.33, and 293.67
C) 723, 748, and 617
D) 361.50, 374.00, and 308.50

15. With which of the following research questions could you use a two-way ANOVA model for the analysis?

A) Is your score on the midterm exam a good predictor of your score on the final exam?
B) Do your level of stress (high or low) and level of close friendships (several close friends or few close friends) effect the number of days you are sick each year?
C) Does the number of days you exercise per week effect your weight loss?
D) Do your favorite color (red, green, blue, or pink) and your weight (overweight, underweight, or appropriate weight) determine whether your first child will be a boy or girl?

16. In a two-way ANOVA, both factors must have the same number of levels.

TRUE

FALSE

17. Plots that show graphs of the marginal means of each factor at every level are called _____.

A) profile plots
B) box plots
C) contour plots
D) factor plots

18. Two-way ANOVA assumes that the observations are independent, the populations have equal variances, and the response variable is approximately normally distributed within each treatment group.

TRUE

FALSE

19. A study was done to determine whether gender and type of student (graduate students or undergraduate) performed differently on a final exam in a statistics course. The mean scores of each group are shown below.

		Gender	
		Female	**Male**
Type of student	**Undergraduate**	78	81
	Graduate	73	62

The data is stored in *EX19_Score.xlsx* (See D2L)

What is the name of the two-way ANOVA needed for the analysis?

A) 2 × 2 ANOVA
B) 4 × 4 ANOVA
C) 4 × 2 ANOVA

D) 2 × 4 ANOVA

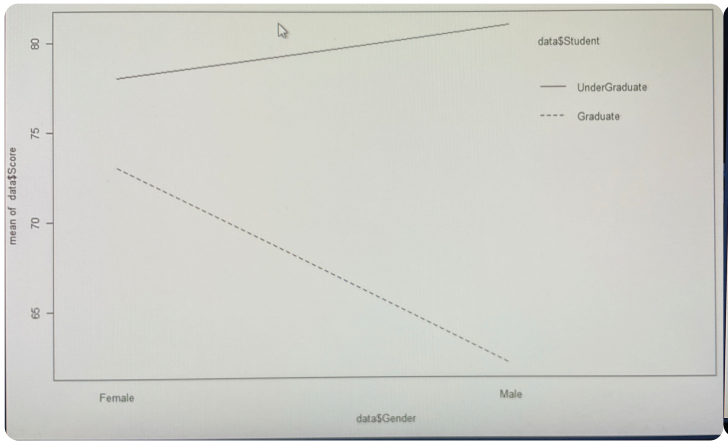
20. Using data from Exercise 19 create an interaction plot.

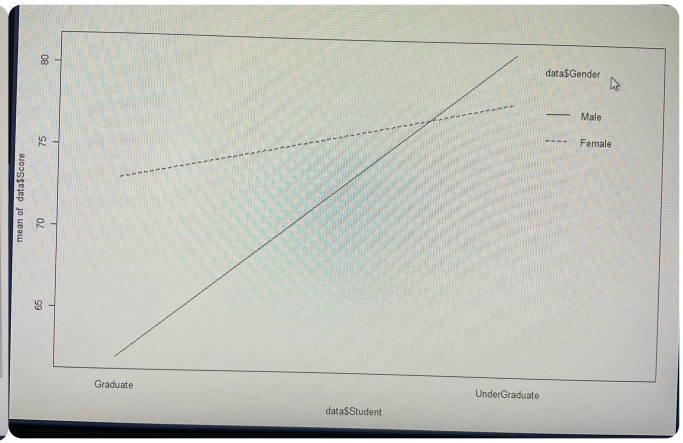
I encourage you to construct interaction plots in R as follows

```
interaction.plot(AvgScore$Student, AvgScore$Gender, AvgScore$Score)

interaction.plot(AvgScore$Gender, AvgScore$Student, AvgScore$Score)
```

COPY AND PASTE THE PLOTS IN THE SPACE BELOW





Does there appear to be any interactions?

A) Yes
B) No

21. Based on the profile plot created in exercise 20, which main effect appears to be the largest?

A) Gender
B) Student

22. An experiment is done as a 3 × 5 ANOVA and the experiment is replicated twice. What is the total number of observations in this experiment?

A) 30
B) 15
C) It cannot be determined from the information given.
D) None of the above

23. An experiment is done as a 3 × 5 ANOVA and the experiment is replicated twice. How large should the *F* statistic be to reject the null hypothesis for the first main effect at the .05 level?

A) 3.68
B) 19.43
C) 15
D) None of the above

$df_1 = 2$
$df_2 = 15$

24. An experiment is done as a 3 × 5 ANOVA and the experiment is replicated twice. What are the degrees of freedom of the *F* statistic associated with the second main effect?

A) 4 and 15
B) 3 and 5
C) 4 and 29
D) None of the above

25. An experiment is done as a 3 × 5 ANOVA and the experiment is replicated twice. What are the degrees of freedom of the *F* statistic associated with the interaction effect?

A) 8 and 15
B) 15 and 30
C) 6 and 30
D) None of the above

26. An experiment is done as a 3 × 5 ANOVA and the experiment is replicated twice. How large should the *F* statistic be to reject the null hypothesis for the interaction at the .05 level?

A) 1
B) 3.22
C) 2.64
D) None of the above

27. A study was done to determine how gender and stress level affect one's self-rated level of happiness on a scale of 1 to 10. The data are shown below. (Use the Excel file ***EX27_Happiness.xls***)

Happiness	Gender	Stress
6	Female	High
7	Female	Low
8	Male	High
10	Male	Low
5	Female	High
7	Female	Low
5	Male	High
10	Male	Low

Running a two-way ANOVA in using your own R script or importing the data in Rcmdr:

Note: To begin with include the interaction term in your ANOVA model and assume all assumptions have been reasonably met.

What is the *P*-value for the **main effect for gender**?

A) Less than or equal to .01

B) Between .02 and .05
C) Between .06 and .08
D) Greater than .09

28. A study was done to determine how gender and stress level affect one's self-rated level of happiness on a scale of 1 to 10. The data are shown in exercise 27.

What is the *P*-value for the main effect for stress? (*Note:* Include the interaction in your ANOVA model and assume all assumptions have been reasonably met.)

A) Less than or equal to .01
B) Between .02 and .05
C) Between .06 and .08
D) Greater than .09

29. A study was done to determine how gender and stress level affect one's self-rated level of happiness on a scale of 1 to 10. The data are shown in exercise 27.

What is the *P*-value for the interaction effect? (*Note:* Assume all assumptions have been reasonably met.)
A) Less than or equal to .01
B) Between .02 and .05
C) Between .06 and .08
D) Greater than .09

30. A research project studied the physical properties of wood materials constructed by bonding together small flakes of wood. The two factors considered were the size of the flakes and species of tree. The sizes of the flakes were S_1: 0.15 inches by 2 inches and S_2: 0.25 inches by 2 inches, and the species of tree used were aspen, birch, and maple. For each combination of flake size and tree species, three samples of wood material were constructed.

For each sample, the physical property measured was the tension modulus of elasticity in the direction perpendicular to the alignment of the flakes, in pounds per square inch (psi)

The table below gives the means for the different size-species groups:

		Species		
		Aspen	**Birch**	**Maple**
Size	**S1**	387.33	292.67	323.33
	S2	335.67	455.33	293.67

For these data, a two-way ANOVA was run and the partial ANOVA table is given below.

Source	**Sum of squares**	**DF**	**Mean square**	***F***
Size	157.44	1	157.44	0.0128
Species	14 512	2	7,256	0.5917
Size*Species	41,707	2	20,853.5	1.7

Source	Sum of squares	DF	Mean square	*F*
Error	147,144	12	12,262	*
Total	**203,520**	**17**	x	*

What is the value of the sum of squares for error, SSE?

A) 1021.8
B) 14,512.0
C) 20,853.5
D) 147,144.0

31. What is the value of the degrees of freedom for the model, DFM?

A) 1
B) 2
C) 5
D) 6

32. What is the value of the *F* statistic for testing interaction?

A) 0.59
B) 1.70
C) 3.40
D) 202.33