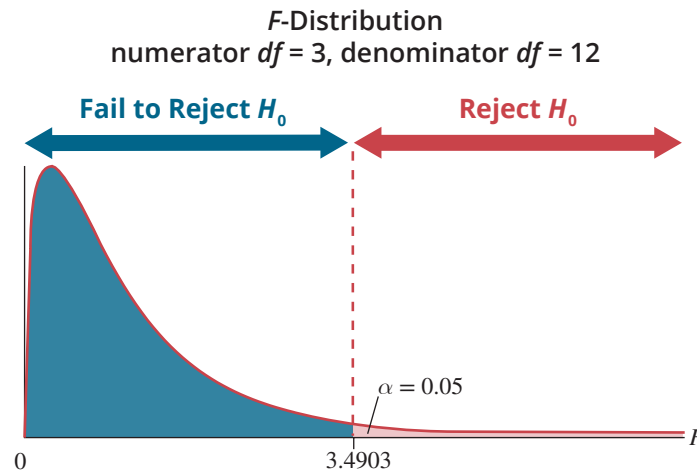


From Figure 15.4.4, the calculated value of the test statistic is 13.4435 and the  $P$ -value is 0.00038. Figure 15.4.7 shows the rejection region. Since 13.4435 is larger than 3.4903 and the  $P$ -value is less than  $\alpha = 0.05$ , we reject the null hypothesis that age has no effect on average salary. There is persuasive evidence at  $\alpha = 0.05$  that the mean salaries are significantly affected by age.



**Figure 15.4.7**

When using the two-way ANOVA, it is important to remember that the assumptions of normality, equal variances, and independent random samples should be satisfied in order for the test to produce meaningful results. As noted previously, the ANOVA test is robust with regards to the normality and equal variance assumptions.

An important concept in statistics related to ANOVA is the **design of experiments** (DOE). It is much easier to analyze a properly designed experiment as opposed to a poorly designed one. If you are planning on collecting data, you should involve a statistician at the onset, not just employ them to do the analysis once the data has been collected. Experimental design allows you to estimate the effects of several variables simultaneously, thus resulting in a more efficient collection of data that will be much easier to analyze. Design of experiments is a very extensive field of study for which entire books and courses have been developed, but is beyond the scope of this book.

## 15.4 Exercises

### Basic Concepts

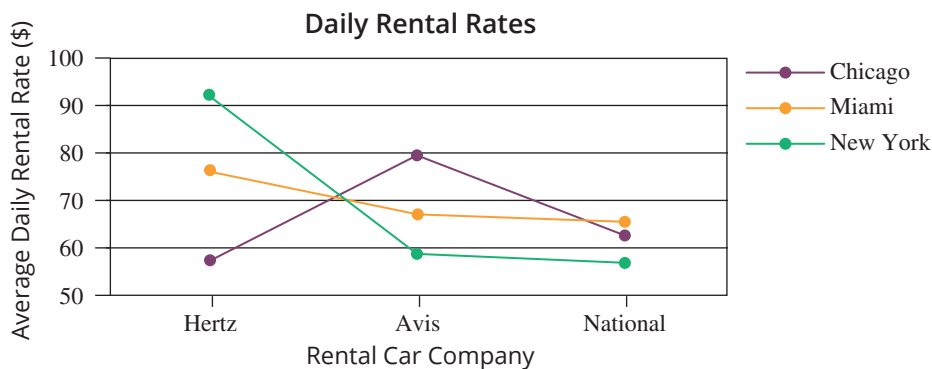
1. What is the difference between a randomized block design and a factorial design?
2. What is a complete factorial experiment?
3. What is an interaction plot? What kind of information does this plot give us?
4. Why is it important to determine if there is interaction between the two variables of interest in a factorial design?
5. Is it possible to perform a two-way analysis of variance if interaction exists between the two variables of interest? Explain why or why not.
6. Identify the four components that make up the total sum of squares in a complete factorial model.

7. Give the degrees of freedom associated with each component of the total sum of squares.
8. What is the test statistic for a test of interaction between factors? What are the degrees of freedom associated with this test statistic?
9. If there is enough evidence to reject the null hypothesis in a test for interaction, may we proceed with the main effects tests? Explain.
10. What is the test statistic for the main effects test for Factor A? What are the degrees of freedom associated with this test statistic?
11. What is the test statistic for the main effects test for Factor B? What are the degrees of freedom associated with this test statistic?
12. What are the rejection rules for the main effects tests? Can  $P$ -values be used as rejection criteria?

## Exercises

13. The following table contains the results of a survey of daily rental rates of a mid-size car for three major rental car companies at three airport locations on three different days during the year.

Daily Rental Rates of Mid-Size Cars (\$)			
	New York	Chicago	Miami
Hertz	93.99	54.99	71.99
	90.99	63.99	87.99
	96.99	57.99	68.99
Avis	58.86	81.99	61.99
	52.10	85.99	70.99
	68.98	71.99	66.99
National	56.00	64.99	66.00
	63.00	67.00	58.99
	52.00	52.99	71.99



- a. Identify the dependent variable. Identify the two factors. In the associated computer output, which variable is Factor A and which variable is Factor B?
- b. Consider the graph of the average daily rental rates for each of the major car rental companies by airport location. Does there appear to be any interaction between the variables airport location and major car rental company?

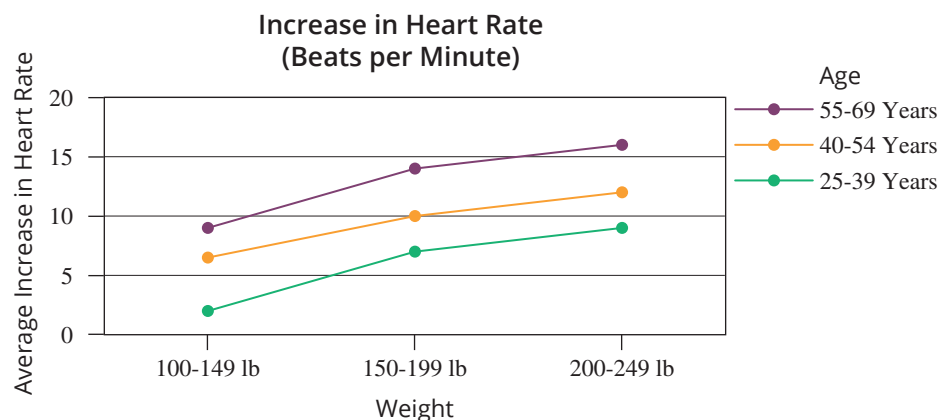
- c. The results of the two-way ANOVA for the study are given in the following table. Perform a hypothesis test to determine if there is any interaction between the variables major rental car company and airport location at  $\alpha = 0.05$ . Does this agree with your observation in part b.?

ANOVA			
Source of Variation	SS	df	MS
Sample	1011.7730	2	505.8865
Columns	58.7126	2	29.3563
Interaction	2514.3099	4	628.5775
Within	819.1289	18	45.5072
Total	4403.9244	26	

- d. If there is no interaction found in part c., is there sufficient evidence to conclude that there is a significant difference among the average daily rental rates for mid-size cars for the three rental car companies at the 0.05 level?
14. A doctor is interested in determining the increase in average heart rate caused by a medication used for treating high blood pressure. The doctor believes that the increase in heart rate will be related to two factors: the age of a person and the weight of a person. To test this theory, the doctor randomly selects two patients in each of the age and weight categories listed in the following table and determines the increase in heart rate (in beats per minute) of each patient 15 minutes after administering the drug. The results of the study are as follows.

Increase in Heart Rate (Beats per Minute)			
	25-39 Years	40-54 Years	55-69 Years
	2	7	11
<b>100-149 Pounds</b>	2	6	7
	7	11	16
<b>150-199 Pounds</b>	7	9	12
	10	13	18
<b>200-249 Pounds</b>	8	11	14

- a. Identify the dependent variable. Identify the two factors. In the associated computer output, which variable is Factor A and which variable is Factor B?
- b. Consider the following graph of the average increase in heart rate for each of the weight and age categories. Does there appear to be any interaction between the age and weight variables? Explain.



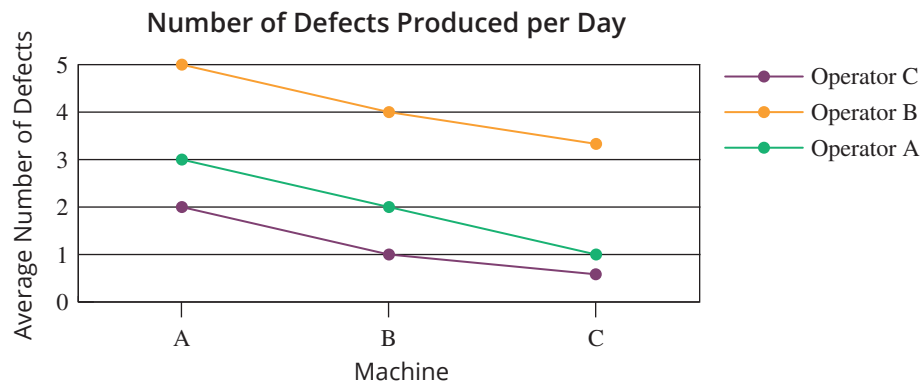
- c. The results of the two-way ANOVA for the study are given in the following table. Perform a hypothesis test to determine if there is any interaction between the variables age and weight at  $\alpha = 0.01$ . Does this agree with your observation in part b.?

ANOVA			
Source of Variation	SS	df	MS
Sample	133.0000	2	66.5000
Columns	147.0000	2	73.5000
Interaction	2.0000	4	0.5000
Within	30.5000	9	3.3889
Total	312.5000	17	

- d. Is there sufficient evidence to conclude that there is a significant difference among the average increases in heart rate for the different weight categories? Use  $\alpha = 0.01$ .
- e. Is there sufficient evidence to conclude that there is a significant difference among the average increases in heart rate for the different age groups? Use  $\alpha = 0.01$ .
15. A supervisor of a manufacturing plant is interested in relating the average number of defects produced per day to two factors: the operator working the machine and the machine itself. The supervisor randomly assigns each operator to use each machine for three days and records the number of defects produced per day. The results of the study are as follows.

Number of Defects Produced per Day			
	Operator A	Operator B	Operator C
Machine A	3	7	3
	3	5	2
	3	3	1
Machine B	2	6	2
	2	4	1
	2	2	0
Machine C	1	5	1
	1	3	0
	1	2	1

- a. Identify the dependent variable. Identify the two factors. In the associated computer output, which variable is Factor A and which variable is Factor B?
- b. Consider the following graph of the average number of defects produced per day for each of the operators by machine. Does there appear to be any interaction between the variables operator and machine?



- c. The results of the two-way ANOVA for the supervisor's survey of the number of defects produced per day are given in the following table. Perform a hypothesis test to determine if there is any interaction between the machine and operator variables. Use  $\alpha = 0.10$ . Does this agree with your observation in part b.?

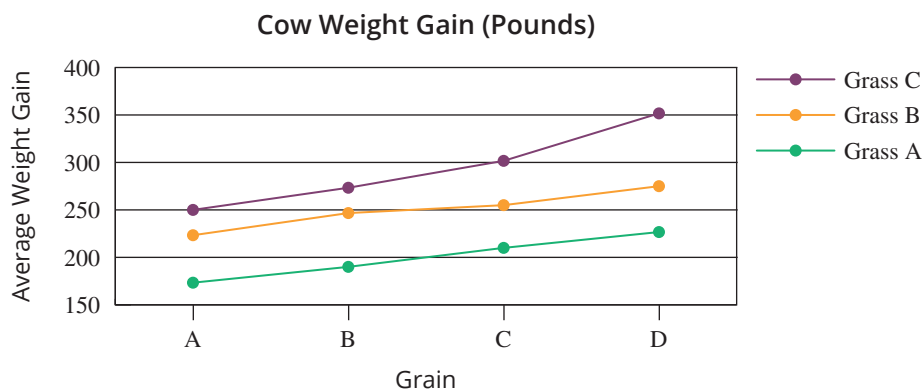
ANOVA			
Source of Variation	SS	df	MS
Sample	12.6667	2	6.3333
Columns	40.2222	2	20.1111
Interaction	0.4444	4	0.1111
Within	25.3333	18	1.4074
Total	78.6667	26	

- d. Is there sufficient evidence to conclude that there is a significant difference among the average number of defects produced per day for the different machines? Use  $\alpha = 0.10$ .
- e. Is there sufficient evidence to conclude that there is a significant difference among the average number of defects produced per day for the different operators? Use  $\alpha = 0.10$ .
16. A dairy farmer thinks that the average weight gain of his cows depends on two factors: the type of grain that they are fed and the type of grass that they are fed. The dairy farmer has four different types of grain from which to choose and three different types of grass from which to choose. He would like to determine if there is a particular combination of grain and grass that would lead to the greatest weight gain on average for his cows. He randomly selects three one-year-old cows and assigns them to each of the possible combinations of grain and grass. After one year he records the weight gain for each cow (in pounds) with the following results.

Cow Weight Gain (Pounds)			
	Grass A	Grass B	Grass C
Grain A	175	225	250
	160	215	240
	185	230	260

<b>Grain B</b>	190	245	275
	185	240	260
	195	255	285
<b>Grain C</b>	210	255	300
	200	245	310
	220	265	295
<b>Grain D</b>	225	275	350
	235	270	360
	220	280	345

- Identify the dependent variable. Identify the two factors. In the associated computer output, which variable is Factor A and which variable is Factor B?
- Consider the following graph of the average weight gain of the cows for each of the possible combinations of grass and grain. Does there appear to be any interaction between the grass and grain variables?



- The results of the two-way ANOVA for the farmer's study are given in the following table. Perform a hypothesis test to determine if there is any interaction between the variables grass and grain at  $\alpha = 0.05$ . Does this agree with your observation in part **b**?

ANOVA			
Source of Variation	SS	df	MS
Sample	23097.2222	3	7699.0741
Columns	53272.2222	2	26636.1111
Interaction	3127.7778	6	521.2963
Within	1916.6667	24	79.8611
Total	81413.8889	35	

- If there is no interaction found in part **c**., is there sufficient evidence to conclude that there is a significant difference in the average weight gains among the cows for the four different types of grain? Use  $\alpha = 0.05$ .
- Is there sufficient evidence to conclude that there is a significant difference in the average weight gains among the cows for the three different types of grass? Use  $\alpha = 0.05$ .

17. The partially completed analysis of variance table given below is taken from the article, “Power and Status, Exchange, Attribution, and Expectation States (Small Group Research).”<sup>5</sup> The experimenters investigated the effects of power and knowledge on one’s emotional reaction in a study involving 52 students selected from a large private university. Each of the factors was run at two levels, with 13 subjects at each of the four different factor combinations.

ANOVA				
<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>
Power	1.2700	1		
Knowledge	0.2500	1		
Interaction		1		
Error	4.1400	48		
Total	5.6700	51		

- Identify the dependent variable. Identify the two factors. In the associated computer output, which variable is Factor A and which variable is Factor B?
- Complete the ANOVA table.
- Can we conclude, with  $\alpha = 0.10$ , that there is interaction between power and knowledge?
- With  $\alpha = 0.05$ , can we conclude that there is a significant difference in the two levels of power?
- With  $\alpha = 0.05$ , can we conclude that there is a significant difference in the two levels of knowledge?