Chapter 13: *F* DISTRIBUTION AND ONE-WAY ANOVA

|  |  |
| --- | --- |
| Exercise 1. | *There are five basic assumptions that must be fulfilled in order to perform a One-Way ANOVA test. What are they?*  *Write one assumption.* |
| Solution | Each population from which a sample is taken is assumed to be normal. |
| Exercise 2. | *There are five basic assumptions that must be fulfilled in order to perform a One-Way ANOVA test. What are they?*  *Write another assumption.* |
| Solution | All samples are randomly selected and independent. |
| Exercise 3. | *There are five basic assumptions that must be fulfilled in order to perform a One-Way ANOVA test. What are they?*  *Write a third assumption.* |
| Solution | The populations are assumed to have equal standard deviations (or variances). |
| Exercise 4. | *There are five basic assumptions that must be fulfilled in order to perform a One-Way ANOVA test. What are they?*  *Write a fourth assumption.* |
| Solution | The factor is a categorical variable. |
| Exercise 5. | *There are five basic assumptions that must be fulfilled in order to perform a One-Way ANOVA test. What are they?*  *Write the final assumption.* |
| Solution | The response is a numerical value. |
| Exercise 6. | *State the null hypothesis for a One-Way ANOVA test if there are four groups.* |
| Solution | *H*0: *μ*1 = *μ*2 = *μ*3 = *μ*4 |
| Exercise 7. | *State the alternative hypothesis for a one-way ANOVA test if there are three groups.* |
| Solution | *Ha:* At least two of the group means μ1, μ2, μ3 are not equal. |
| Exercise 8. | *When do you use an ANOVA test?* |
| Solution | Use ANOVA when you want to test if the means of three or more populations are equal. |
| Exercise 9. | *Groups of men from three different areas of the country are to be tested for mean weight. The entries in the table are the weights for the different groups. The One-Way ANOVA results are shown in Table 13.13.*   |  |  |  | | --- | --- | --- | | ***Group 1*** | ***Group 2*** | ***Group 3*** | | *216* | *202* | *170* | | *198* | *213* | *165* | | *240* | *284* | *182* | | *187* | *228* | *197* | | *176* | *210* | *201* |   *Table 13.13*  *What is the Sum of Squares Factor?* |
| Solution | 4,939.2 |
| Exercise 10. | *Groups of men from three different areas of the country are to be tested for mean weight. The entries in the table are the weights for the different groups. The One-Way ANOVA results are shown in Table 13.13.*   |  |  |  | | --- | --- | --- | | ***Group 1*** | ***Group 2*** | ***Group 3*** | | *216* | *202* | *170* | | *198* | *213* | *165* | | *240* | *284* | *182* | | *187* | *228* | *197* | | *176* | *210* | *201* |   *Table 13.13*  *What is the Sum of Squares Error?* |
| Solution | 7,920.4 |
| Exercise 11. | *Groups of men from three different areas of the country are to be tested for mean weight. Groups of men from three different areas of the country are to be tested for mean weight. The entries in the table are the weights for the different groups. The One-Way ANOVA results are shown in Table 13.13.*   |  |  |  | | --- | --- | --- | | ***Group 1*** | ***Group 2*** | ***Group 3*** | | *216* | *202* | *170* | | *198* | *213* | *165* | | *240* | *284* | *182* | | *187* | *228* | *197* | | *176* | *210* | *201* |   *Table 13.13*  *What is the df for the numerator?* |
| Solution | 2 |
| Exercise 12. | *Groups of men from three different areas of the country are to be tested for mean weight. The entries in the table are the weights for the different groups. The One-Way ANOVA results are shown in Table 13.13.*   |  |  |  | | --- | --- | --- | | ***Group 1*** | ***Group 2*** | ***Group 3*** | | *216* | *202* | *170* | | *198* | *213* | *165* | | *240* | *284* | *182* | | *187* | *228* | *197* | | *176* | *210* | *201* |   *Table 13.13*  *What is the df for the denominator?* |
| Solution | 12 |
| Exercise 13. | *Groups of men from three different areas of the country are to be tested for mean weight. The entries in the table are the weights for the different groups. The One-Way ANOVA results are shown in Table 13.13.*   |  |  |  | | --- | --- | --- | | ***Group 1*** | ***Group 2*** | ***Group 3*** | | *216* | *202* | *170* | | *198* | *213* | *165* | | *240* | *284* | *182* | | *187* | *228* | *197* | | *176* | *210* | *201* |   *Table 13.13*  *What is the Mean Square Factor?* |
| Solution | 2,469.6 |
| Exercise 14. | *Groups of men from three different areas of the country are to be tested for mean weight. The entries in the table are the weights for the different groups. The One-Way ANOVA results are shown in Table 13.13.*   |  |  |  | | --- | --- | --- | | ***Group 1*** | ***Group 2*** | ***Group 3*** | | *216* | *202* | *170* | | *198* | *213* | *165* | | *240* | *284* | *182* | | *187* | *228* | *197* | | *176* | *210* | *201* |   *Table 13.13*  *What is the Mean Square Error?* |
| Solution | 660.03 |
| Exercise 15. | *Groups of men from three different areas of the country are to be tested for mean weight. The entries in the table are the weights for the different groups. The One-Way ANOVA results are shown in Table 13.13.*   |  |  |  | | --- | --- | --- | | ***Group 1*** | ***Group 2*** | ***Group 3*** | | *216* | *202* | *170* | | *198* | *213* | *165* | | *240* | *284* | *182* | | *187* | *228* | *197* | | *176* | *210* | *201* |   *Table 13.13*  *What is the F statistic?* |
| Solution | 3.7416 |
| Exercise 16. | *Girls from four different soccer teams are to be tested for mean goals scored per game. The entries in the table are the goals per game for the different teams. The one-way ANOVA results are shown in Table 13.14.*   |  |  |  |  | | --- | --- | --- | --- | | ***Team 1*** | ***Team 2*** | ***Team 3*** | ***Team 4*** | | *1* | *2* | *0* | *3* | | *2* | *3* | *1* | *4* | | *0* | *2* | *1* | *4* | | *3* | *4* | *0* | *3* | | *2* | *4* | *0* | *2* |   *Table 13.14*  *What is SSbetween?* |
| Solution | 25.75 |
| Exercise 17. | *Girls from four different soccer teams are to be tested for mean goals scored per game. The entries in the table are the goals per game for the different teams. The one-way ANOVA results are shown in Table 13.14.*   |  |  |  |  | | --- | --- | --- | --- | | ***Team 1*** | ***Team 2*** | ***Team 3*** | ***Team 4*** | | *1* | *2* | *0* | *3* | | *2* | *3* | *1* | *4* | | *0* | *2* | *1* | *4* | | *3* | *4* | *0* | *3* | | *2* | *4* | *0* | *2* |   *Table 13.14*  *What is the df for the numerator?* |
| Solution | 3 |
| Exercise 18. | *Girls from four different soccer teams are to be tested for mean goals scored per game. The entries in the table are the goals per game for the different teams. The one-way ANOVA results are shown in Table 13.14.*   |  |  |  |  | | --- | --- | --- | --- | | ***Team 1*** | ***Team 2*** | ***Team 3*** | ***Team 4*** | | *1* | *2* | *0* | *3* | | *2* | *3* | *1* | *4* | | *0* | *2* | *1* | *4* | | *3* | *4* | *0* | *3* | | *2* | *4* | *0* | *2* |   *Table 13.14*  *What is MSbetween?* |
| Solution | 8.5833 |
| Exercise 19. | *Girls from four different soccer teams are to be tested for mean goals scored per game. The entries in the table are the goals per game for the different teams. The one-way ANOVA results are shown in Table 13.14.*   |  |  |  |  | | --- | --- | --- | --- | | ***Team 1*** | ***Team 2*** | ***Team 3*** | ***Team 4*** | | *1* | *2* | *0* | *3* | | *2* | *3* | *1* | *4* | | *0* | *2* | *1* | *4* | | *3* | *4* | *0* | *3* | | *2* | *4* | *0* | *2* |   *Table 13.14*  *What is SSwithin?* |
| Solution | 13.2 |
| Exercise 20. | *Girls from four different soccer teams are to be tested for mean goals scored per game. The entries in the table are the goals per game for the different teams. The one-way ANOVA results are shown in Table 13.14.*   |  |  |  |  | | --- | --- | --- | --- | | ***Team 1*** | ***Team 2*** | ***Team 3*** | ***Team 4*** | | *1* | *2* | *0* | *3* | | *2* | *3* | *1* | *4* | | *0* | *2* | *1* | *4* | | *3* | *4* | *0* | *3* | | *2* | *4* | *0* | *2* |   *Table 13.14*  *What is the df for the denominator?* |
| Solution | 16 |
| Exercise 21. | *Girls from four different soccer teams are to be tested for mean goals scored per game. The entries in the table are the goals per game for the different teams. The one-way ANOVA results are shown in Table 13.14.*   |  |  |  |  | | --- | --- | --- | --- | | ***Team 1*** | ***Team 2*** | ***Team 3*** | ***Team 4*** | | *1* | *2* | *0* | *3* | | *2* | *3* | *1* | *4* | | *0* | *2* | *1* | *4* | | *3* | *4* | *0* | *3* | | *2* | *4* | *0* | *2* |   *Table 13.14*  *What is MSwithin?* |
| Solution | 0.825 |
| Exercise 22. | *Girls from four different soccer teams are to be tested for mean goals scored per game. The entries in the table are the goals per game for the different teams. The one-way ANOVA results are shown in Table 13.14.*   |  |  |  |  | | --- | --- | --- | --- | | ***Team 1*** | ***Team 2*** | ***Team 3*** | ***Team 4*** | | *1* | *2* | *0* | *3* | | *2* | *3* | *1* | *4* | | *0* | *2* | *1* | *4* | | *3* | *4* | *0* | *3* | | *2* | *4* | *0* | *2* |   *Table 13.14*  *What is the F statistic?* |
| Solution | 10.40 |
| Exercise 23. | *Girls from four different soccer teams are to be tested for mean goals scored per game. The entries in the table are the goals per game for the different teams. The one-way ANOVA results are shown in Table 13.14.*   |  |  |  |  | | --- | --- | --- | --- | | ***Team 1*** | ***Team 2*** | ***Team 3*** | ***Team 4*** | | *1* | *2* | *0* | *3* | | *2* | *3* | *1* | *4* | | *0* | *2* | *1* | *4* | | *3* | *4* | *0* | *3* | | *2* | *4* | *0* | *2* |   *Table 13.14*  *Judging by the F statistic, do you think it is likely or unlikely that you will reject the null hypothesis?* |
| Solution | Because a one-way ANOVA test is always right tailed, a high *F* statistic corresponds to a low *p*-value, so it is likely that we will reject the null hypothesis. |
| Exercise 24. | *An F statistic can have what values?* |
| Solution | greater than or equal to zero |
| Exercise 25. | *What happens to the curves as the degrees of freedom for the numerator and the denominator get larger?* |
| Solution | The curves approximate the normal distribution. |
| Exercise 26. | *Four basketball teams took a random sample of players regarding how high each player can jump (in inches). The results are in Table 13.15:*   |  |  |  |  |  | | --- | --- | --- | --- | --- | | *Team 1* | *Team 2* | *Team 3* | *Team 4* | *Team 5* | | *36* | *32* | *48* | *38* | *41* | | *42* | *35* | *50* | *44* | *39* | | *51* | *38* | *39* | *46* | *40* |   *Table 13.15*  *What is the df(num)?* |
| Solution | four |
| Exercise 27. | *Four basketball teams took a random sample of players regarding how high each player can jump (in inches). The results are in Table 13.15:*   |  |  |  |  |  | | --- | --- | --- | --- | --- | | *Team 1* | *Team 2* | *Team 3* | *Team 4* | *Team 5* | | *36* | *32* | *48* | *38* | *41* | | *42* | *35* | *50* | *44* | *39* | | *51* | *38* | *39* | *46* | *40* |   *Table 13.15*  *What is the df(denom)?* |
| Solution | ten |
| Exercise 28. | *Four basketball teams took a random sample of players regarding how high each player can jump (in inches). The results are in Table 13.15:*   |  |  |  |  |  | | --- | --- | --- | --- | --- | | *Team 1* | *Team 2* | *Team 3* | *Team 4* | *Team 5* | | *36* | *32* | *48* | *38* | *41* | | *42* | *35* | *50* | *44* | *39* | | *51* | *38* | *39* | *46* | *40* |   *Table 13.15*  *What are the Sum of Squares and Mean Square Factor?* |
| Solution | *SS* = 195.6; *MS* = 48.9 |
| Exercise 29. | *Four basketball teams took a random sample of players regarding how high each player can jump (in inches). The results are in Table 13.15:*   |  |  |  |  |  | | --- | --- | --- | --- | --- | | *Team 1* | *Team 2* | *Team 3* | *Team 4* | *Team 5* | | *36* | *32* | *48* | *38* | *41* | | *42* | *35* | *50* | *44* | *39* | | *51* | *38* | *39* | *46* | *40* |   *Table 13.15*  *What are the Sum of Squares and Mean Square Errors?* |
| Solution | Solution *SS* = 237.33; *MS* = 23.73 |
| Exercise 30. | *Four basketball teams took a random sample of players regarding how high each player can jump (in inches). The results are in Table 13.15:*   |  |  |  |  |  | | --- | --- | --- | --- | --- | | *Team 1* | *Team 2* | *Team 3* | *Team 4* | *Team 5* | | *36* | *32* | *48* | *38* | *41* | | *42* | *35* | *50* | *44* | *39* | | *51* | *38* | *39* | *46* | *40* |   *Table 13.15*  *What is the F statistic?* |
| Solution | 2.06 |
| Exercise 31. | *Four basketball teams took a random sample of players regarding how high each player can jump (in inches). The results are in Table 13.15:*   |  |  |  |  |  | | --- | --- | --- | --- | --- | | *Team 1* | *Team 2* | *Team 3* | *Team 4* | *Team 5* | | *36* | *32* | *48* | *38* | *41* | | *42* | *35* | *50* | *44* | *39* | | *51* | *38* | *39* | *46* | *40* |   *Table 13.15*  *What is the p-value?* |
| Solution | 0.1614 |
| Exercise 32. | *Four basketball teams took a random sample of players regarding how high each player can jump (in inches). The results are in Table 13.15:*   |  |  |  |  |  | | --- | --- | --- | --- | --- | | *Team 1* | *Team 2* | *Team 3* | *Team 4* | *Team 5* | | *36* | *32* | *48* | *38* | *41* | | *42* | *35* | *50* | *44* | *39* | | *51* | *38* | *39* | *46* | *40* |   *Table 13.15*  *At the 5% significance level, is there a difference in the mean jump heights among the teams?* |
| Solution | No, there is not enough evidence to show that there is a significant difference in the heights of the jumps of the players among the teams. |
| Exercise 33. | *A video game developer is testing a new game on three different groups. Each group represents a different target market for the game. The developer collects scores from a random sample from each group. The results are shown Table 13.16:*   |  |  |  | | --- | --- | --- | | ***Group A*** | ***Group B*** | ***Group C*** | | *101* | *151* | *101* | | *108* | *149* | *109* | | *98* | *160* | *198* | | *107* | *112* | *186* | | *111* | *126* | *160* |   *Table 13.16*  *What is the df(num)?* |
| Solution | two |
| Exercise 34. | *A video game developer is testing a new game on three different groups. Each group represents a different target market for the game. The developer collects scores from a random sample from each group. The results are shown Table 13.16:*   |  |  |  | | --- | --- | --- | | ***Group A*** | ***Group B*** | ***Group C*** | | *101* | *151* | *101* | | *108* | *149* | *109* | | *98* | *160* | *198* | | *107* | *112* | *186* | | *111* | *126* | *160* |   *Table 13.16*  *What is the df(denom)?* |
| Solution | 12 |
| Exercise 35. | *A video game developer is testing a new game on three different groups. Each group represents a different target market for the game. The developer collects scores from a random sample from each group. The results are shown Table 13.16:*   |  |  |  | | --- | --- | --- | | ***Group A*** | ***Group B*** | ***Group C*** | | *101* | *151* | *101* | | *108* | *149* | *109* | | *98* | *160* | *198* | | *107* | *112* | *186* | | *111* | *126* | *160* |   *Table 13.16*  *What are the SSbetween and MSbetween?* |
| Solution | *SS* = 5,700.4  *MS* = 2,850.2 |
| Exercise 36. | *A video game developer is testing a new game on three different groups. Each group represents a different target market for the game. The developer collects scores from a random sample from each group. The results are shown Table 13.16:*   |  |  |  | | --- | --- | --- | | ***Group A*** | ***Group B*** | ***Group C*** | | *101* | *151* | *101* | | *108* | *149* | *109* | | *98* | *160* | *198* | | *107* | *112* | *186* | | *111* | *126* | *160* |   *Table 13.16*  *What are the SSwithin and MSwithin?* |
| Solution | *SS* = 9,474  *MS* = 789.5 |
| Exercise 37. | *A video game developer is testing a new game on three different groups. Each group represents a different target market for the game. The developer collects scores from a random sample from each group. The results are shown Table 13.16:*   |  |  |  | | --- | --- | --- | | ***Group A*** | ***Group B*** | ***Group C*** | | *101* | *151* | *101* | | *108* | *149* | *109* | | *98* | *160* | *198* | | *107* | *112* | *186* | | *111* | *126* | *160* |   *Table 13.16*  *What is the F statistic?* |
| Solution | 3.6101 |
| Exercise 38. | *A video game developer is testing a new game on three different groups. Each group represents a different target market for the game. The developer collects scores from a random sample from each group. The results are shown Table 13.16:*   |  |  |  | | --- | --- | --- | | ***Group A*** | ***Group B*** | ***Group C*** | | *101* | *151* | *101* | | *108* | *149* | *109* | | *98* | *160* | *198* | | *107* | *112* | *186* | | *111* | *126* | *160* |   *Table 13.16*  *What is the p-value?* |
| Solution | 0.0592 |
| Exercise 39. | *A video game developer is testing a new game on three different groups. Each group represents a different target market for the game. The developer collects scores from a random sample from each group. The results are shown Table 13.16:*   |  |  |  | | --- | --- | --- | | ***Group A*** | ***Group B*** | ***Group C*** | | *101* | *151* | *101* | | *108* | *149* | *109* | | *98* | *160* | *198* | | *107* | *112* | *186* | | *111* | *126* | *160* |   *Table 13.16*  *At the 10% significance level, are the scores among the different groups different?* |
| Solution | Yes, there is enough evidence to show that the scores among the groups are statistically significant at the 10% level. |
| Exercise 40. | *Suppose a group is interested in determining whether teenagers obtain their drivers licenses at approximately the same average age across the country. Suppose that the following data are randomly collected from five teenagers in each region of the country. The numbers represent the age at which teenagers obtained their drivers licenses.*   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | **Northeast** | **South** | **West** | **Central** | **East** | |  | 16.3 | 16.9 | 16.4 | 16.2 | 17.1 | |  | 16.1 | 16.5 | 16.5 | 16.6 | 17.2 | |  | 16.4 | 16.4 | 16.6 | 16.5 | 16.6 | |  | 16.5 | 16.2 | 16.1 | 16.4 | 16.8 | | = | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | | *s*2= | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ |   *Table 13.17*  *p-value = \_\_\_\_\_\_* |
| Solution | 0.017 |
| Exercise 41. | *Suppose a group is interested in determining whether teenagers obtain their drivers licenses at approximately the same average age across the country. Suppose that the following data are randomly collected from five teenagers in each region of the country. The numbers represent the age at which teenagers obtained their drivers licenses.*   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | **Northeast** | **South** | **West** | **Central** | **East** | |  | 16.3 | 16.9 | 16.4 | 16.2 | 17.1 | |  | 16.1 | 16.5 | 16.5 | 16.6 | 17.2 | |  | 16.4 | 16.4 | 16.6 | 16.5 | 16.6 | |  | 16.5 | 16.2 | 16.1 | 16.4 | 16.8 | | = | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | | *s*2= | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ |   *State the decisions and conclusion (in complete sentences)for the following preconceived levels of α.*  *α = 0.05*  *a. Decision: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_*  *b. Conclusion: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_* |
| Solution | = 0.05: Reject the null hypothesis; Conclusion: The mean ages teenagers get their drivers licenses are different. |
| Exercise 42. | *Suppose a group is interested in determining whether teenagers obtain their drivers licenses at approximately the same average age across the country. Suppose that the following data are randomly collected from five teenagers in each region of the country. The numbers represent the age at which teenagers obtained their drivers licenses.*   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | **Northeast** | **South** | **West** | **Central** | **East** | |  | 16.3 | 16.9 | 16.4 | 16.2 | 17.1 | |  | 16.1 | 16.5 | 16.5 | 16.6 | 17.2 | |  | 16.4 | 16.4 | 16.6 | 16.5 | 16.6 | |  | 16.5 | 16.2 | 16.1 | 16.4 | 16.8 | | = | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | | *s*2= | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ |   *State the decisions and conclusion (in complete sentences)for the following preconceived levels of α.*  *α = 0.01*  *a. Decision: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_*  *b. Conclusion: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_* |
| Solution | =0.01: Decline to reject null hypothesis; Conclusion: The mean age teenagers get their drivers licenses are not different. |
| Exercise 43. | *There are two assumptions that must be true in order to perform an F test of two variances.*  *Name one assumption that must be true.* |
| Solution | The populations from which the two samples are drawn are normally distributed. |
| Exercise 44. | *There are two assumptions that must be true in order to perform an F test of two variances.*  *What is the other assumption that must be true?* |
| Solution | The two populations are independent of each other. |
| Exercise 45. | *Two coworkers commute from the same building. They are interested in whether or not there is any variation in the time it takes them to drive to work. They each record their times for 20 commutes. The first worker’s times have a variance of 12.1. The second worker’s times have a variance of 16.9. The first worker thinks that he is more consistent with his commute times and that his commute time is shorter. Test the claim at the 10% level.*  *State the null and alternative hypotheses.* |
| Solution | *H*0: *σ*1 = *σ*2; *H*a: *σ*1 < *σ*2 or *H*0: |
| Exercise 46. | *Two coworkers commute from the same building. They are interested in whether or not there is any variation in the time it takes them to drive to work. They each record their times for 20 commutes. The first worker’s times have a variance of 12.1. The second worker’s times have a variance of 16.9. The first worker thinks that he is more consistent with his commute times and that his commute time is shorter. Test the claim at the 10% level.*  *What is s1 in this problem?* |
| Solution | 3.47 |
| Exercise 47. | *Two coworkers commute from the same building. They are interested in whether or not there is any variation in the time it takes them to drive to work. They each record their times for 20 commutes. The first worker’s times have a variance of 12.1. The second worker’s times have a variance of 16.9. The first worker thinks that he is more consistent with his commute times and that his commute time is shorter. Test the claim at the 10% level.*  *What is s2 in this problem?* |
| Solution | 4.11 |
| Exercise 48. | *Two coworkers commute from the same building. They are interested in whether or not there is any variation in the time it takes them to drive to work. They each record their times for 20 commutes. The first worker’s times have a variance of 12.1. The second worker’s times have a variance of 16.9. The first worker thinks that he is more consistent with his commute times and that his commute time is shorter. Test the claim at the 10% level.*  *What is n?* |
| Solution | 20 |
| Exercise 49. | *Two coworkers commute from the same building. They are interested in whether or not there is any variation in the time it takes them to drive to work. They each record their times for 20 commutes. The first worker’s times have a variance of 12.1. The second worker’s times have a variance of 16.9. The first worker thinks that he is more consistent with his commute times and that his commute time is shorter. Test the claim at the 10% level.*  *What is the F statistic?* |
| Solution | 0.7159 |
| Exercise 50. | *Two coworkers commute from the same building. They are interested in whether or not there is any variation in the time it takes them to drive to work. They each record their times for 20 commutes. The first worker’s times have a variance of 12.1. The second worker’s times have a variance of 16.9. The first worker thinks that he is more consistent with his commute times and that his commute time is shorter. Test the claim at the 10% level.*  *What is the p-value?* |
| Solution | 0.2367 |
| Exercise 51. | *Two coworkers commute from the same building. They are interested in whether or not there is any variation in the time it takes them to drive to work. They each record their times for 20 commutes. The first worker’s times have a variance of 12.1. The second worker’s times have a variance of 16.9. The first worker thinks that he is more consistent with his commute times and that his commute time is shorter. Test the claim at the 10% level.*  *Is the claim accurate?* |
| Solution | *No, at the 10% level of significance, we do not reject the null hypothesis and state that the data do not show that the variation in drive times for the first worker is less than the variation in drive times for the second worker.* |
| Exercise 52. | *Two students are interested in whether or not there is variation in their test scores for math class. There are 15 total math tests they have taken so far. The first student’s grades have a standard deviation of 38.1. The second student’s grades have a standard deviation of 22.5. The second student thinks his scores are lower.*  *State the null and alternative hypotheses.* |
| Solution |  |
| Exercise 53. | *Two students are interested in whether or not there is variation in their test scores for math class. There are 15 total math tests they have taken so far. The first student’s grades have a standard deviation of 38.1. The second student’s grades have a standard deviation of 22.5. The second student thinks his scores are lower.*  *What is the F Statistic?* |
| Solution | 2.8674 |
| Exercise 54. | *Two students are interested in whether or not there is variation in their test scores for math class. There are 15 total math tests they have taken so far. The first student’s grades have a standard deviation of 38.1. The second student’s grades have a standard deviation of 22.5. The second student thinks his scores are lower.*  *What is the p-value?* |
| Solution | 0.0290 |
| Exercise 55. | *Two students are interested in whether or not there is variation in their test scores for math class. There are 15 total math tests they have taken so far. The first student’s grades have a standard deviation of 38.1. The second student’s grades have a standard deviation of 22.5. The second student thinks his scores are lower.*  *At the 5% significance level, do we reject the null hypothesis?* |
| Solution | Reject the null hypothesis. There is enough evidence to say that the variance of the grades for the first student is higher than the variance in the grades for the second student. |
| Exercise 56. | *Two cyclists are comparing the variances of their overall paces going uphill. Each cyclist records his or her speeds going up 35 hills. The first cyclist has a variance of 23.8 and the second cyclist has a variance of 32.1. The cyclists want to see if their variances are the same or different.*  *State the null and alternative hypotheses* |
| Solution |  |
| Exercise 57. | *Two cyclists are comparing the variances of their overall paces going uphill. Each cyclist records his or her speeds going up 35 hills. The first cyclist has a variance of 23.8 and the second cyclist has a variance of 32.1. The cyclists want to see if their variances are the same or different.*  *What is the F Statistic?* |
| Solution | 0.7414 |
| Exercise 58. | *Two cyclists are comparing the variances of their overall paces going uphill. Each cyclist records his or her speeds going up 35 hills. The first cyclist has a variance of 23.8 and the second cyclist has a variance of 32.1. The cyclists want to see if their variances are the same or different.*  *At the 5% significance level, what can we say about the cyclists’ variances?* |
| Solution | We decline to reject the null hypothesis. There is not enough evidence to show that the two cyclists have different variances. |
| Exercise 59. | *Three different traffic routes are tested for mean driving time. The entries in the table are the driving times in minutes on the three different routes. The one-way ANOVA results are shown in Table 13.18****.***   |  |  |  | | --- | --- | --- | | ***Route 1*** | ***Route 2*** | ***Route 3*** | | *30* | *27* | *16* | | *32* | *29* | *41* | | *27* | *28* | *22* | | *35* | *36* | *31* |   *Table 13.18*  *State SSbetween, SSwithin, and the F statistic.* | |
| Solution | SSbetween = 26  SSwithin = 441  *F* = 0.2653 | |
| Exercise 60. | *Suppose a group is interested in determining whether teenagers obtain their drivers licenses at approximately the same average age across the country. Suppose that the following data are randomly collected from five teenagers in each region of the country. The numbers represent the age at which teenagers obtained their drivers licenses.*   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | ***Northeast*** | ***South*** | ***West*** | ***Central*** | ***East*** | |  | *16.3* | *16.9* | *16.4* | *16.2* | *17.1* | |  | *16.1* | *16.5* | *16.5* | *16.6* | *17.2* | |  | *16.4* | *16.4* | *16.6* | *16.5* | *16.6* | |  | *16.5* | *16.2* | *16.1* | *16.4* | *16.8* | |  | *\_\_\_\_\_\_\_\_* | *\_\_\_\_\_\_\_\_* | *\_\_\_\_\_\_\_* | *\_\_\_\_\_\_\_* | *\_\_\_\_\_\_\_\_* | | *s2 =* | *\_\_\_\_\_\_\_\_* | *\_\_\_\_\_\_\_\_* | *\_\_\_\_\_\_\_* | *\_\_\_\_\_\_\_* | *\_\_\_\_\_\_\_\_* |   *Table 13.19*  *State the hypotheses*  *H0: \_\_\_\_\_\_\_\_\_\_\_\_*  *Ha: \_\_\_\_\_\_\_\_\_\_\_\_* | |
| Solution | *H0:* *µ*1=*µ*2=*µ*3=*µ*4=*µ*5  *Hα:* At least any two of the group means 1, µ2, …, µ5 are not equal. | |
| Exercise 61. | *Suppose a group is interested in determining whether teenagers obtain their drivers licenses at approximately the same average age across the country. Suppose that the following data are randomly collected from five teenagers in each region of the country. The numbers represent the age at which teenagers obtained their drivers licenses.*   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | ***Northeast*** | ***South*** | ***West*** | ***Central*** | ***East*** | |  | *16.3* | *16.9* | *16.4* | *16.2* | *17.1* | |  | *16.1* | *16.5* | *16.5* | *16.6* | *17.2* | |  | *16.4* | *16.4* | *16.6* | *16.5* | *16.6* | |  | *16.5* | *16.2* | *16.1* | *16.4* | *16.8* | |  | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_ | | *s2 =* | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_ |   *Table 13.20*  *degrees of freedom – numerator: df(num) =*\_\_\_\_\_\_\_\_\_ | |
| Solution | *df*(*num*) = 4 | |
| Exercise 62. | *Suppose a group is interested in determining whether teenagers obtain their drivers licenses at approximately the same average age across the country. Suppose that the following data are randomly collected from five teenagers in each region of the country. The numbers represent the age at which teenagers obtained their drivers licenses.*   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | ***Northeast*** | ***South*** | ***West*** | ***Central*** | ***East*** | |  | *16.3* | *16.9* | *16.4* | *16.2* | *17.1* | |  | *16.1* | *16.5* | *16.5* | *16.6* | *17.2* | |  | *16.4* | *16.4* | *16.6* | *16.5* | *16.6* | |  | *16.5* | *16.2* | *16.1* | *16.4* | *16.8* | |  | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_ | | *s2 =* | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_ |   *Table 13.20*  *Degrees of freedom - denominator: df(denom) =* \_\_\_\_\_\_\_\_\_ | |
| Solution | *df*(*denom*) = 15 | |
| Exercise 63. | *Suppose a group is interested in determining whether teenagers obtain their drivers licenses at approximately the same average age across the country. Suppose that the following data are randomly collected from five teenagers in each region of the country. The numbers represent the age at which teenagers obtained their drivers licenses.*   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | ***Northeast*** | ***South*** | ***West*** | ***Central*** | ***East*** | |  | *16.3* | *16.9* | *16.4* | *16.2* | *17.1* | |  | *16.1* | *16.5* | *16.5* | *16.6* | *17.2* | |  | *16.4* | *16.4* | *16.6* | *16.5* | *16.6* | |  | *16.5* | *16.2* | *16.1* | *16.4* | *16.8* | |  | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_ | | *s2 =* | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_ |   *Table 13.20*  *F statistic =* \_\_\_\_\_\_\_\_\_ | |
| Solution | *F* = 4.22 | |
| Exercise 64. | *Three students, Linda, Tuan, and Javier, are given five laboratory rats each for a nutritional experiment. Each rat's weight is recorded in grams. Linda feeds her rats Formula A, Tuan feeds his rats Formula B, and Javier feeds his rats Formula C. At the end of a specified time period, each rat is weighed again, and the net gain in grams is recorded. Using a significance level of 10%, test the hypothesis that the three formulas produce the same mean weight gain.*   |  |  |  | | --- | --- | --- | | ***Linda's rats*** | ***Tuan's rats*** | ***Javier's rats*** | | *43.5* | *47.0* | *51.2* | | *39.4* | *40.5* | *40.9* | | *41.3* | *38.9* | *37.9* | | *46.0* | *46.3* | *45.0* | | *38.2* | *44.2* | *48.6* |   Table 13.21 Weights of Student Lab Rats | |
| Solution | a. *H0*: *μL* = *μT* = *μJ*  b. at least any two of the means are different  c. *df*(*num*) = 2; *df*(*denom*) = 12  d. *F* distribution  e. 0.67  f. 0.5305  g. Check student’s solution.  h. Decision: Do not reject null hypothesis; Conclusion: There is insufficient evidence to conclude that the means are different. | |
| Exercise 65. | *A grassroots group opposed to a proposed increase in the gas tax claimed that the increase would hurt working-class people the most, since they commute the farthest to work. Suppose that the group randomly surveyed 24 individuals and asked them their daily one-way commuting mileage. The results are in Table 13.22. Using a 5% significance level, test the hypothesis that the three mean commuting mileages are the same.*   |  |  |  | | --- | --- | --- | | ***working-class*** | ***professional(middle incomes)*** | ***professional (wealthy)*** | | *17.8* | *16.5* | *8.5* | | *26.7* | *17.4* | *6.3* | | *49.4* | *22.0* | *4.6* | | *9.4* | *7.4* | *12.6* | | *65.4* | *9.4* | *11.0* | | *47.1* | *2.1* | *28.6* | | *19.5* | *6.4* | *15.4* | | *51.2* | *13.9* | *9.3* |   *Table 13.22* | |
| Solution | Subscripts: wc = Working-class; mi = Professional (middle-incomes); we = Professional (wealthy)  a. H0: µwc = µmi = µwe  b. At least any two of the means are different  c. *df*(*n*) = 2, *df*(*d*) = 21  d. *F*-distribution  e. *F* = 9.10  f. 0.001  g. Check student’s solution.  h.  i. Alpha: 0.05  ii. Decision: Reject the null hypothesis  iii. Reason for decision: The three mean commuting mileages are not equal.  iv. Conclusion: There is sufficient evidence to conclude that the mean commuting mileage for working class, middle-income professional, and wealthy professional are different. | |
| Exercise 66. | *Examine the seven practice laps from Terri Vogel's Log book. Determine whether the mean lap time is statistically the same for the seven practice laps, or if there is at least one lap that has a different mean time from the others.* | |
| Solution | a. *H0*: *μ*1 = *μ*2 = *μ*3 = *μ*4 = *μ*5 = *μ*6 = *μ*7  b. At least two mean lap times are different.  c. *df*(*num*) = 6; *df*(*denom*) = 98  d. *F* distribution  e. 1.69  f. 0.1319  g. Check student’s solution.  h. Decision: Do not reject null hypothesis; Conclusion: There is insufficient evidence to conclude that the mean lap times are different. | |
| Exercise 67. | *Table 13.23 lists the number of pages in four different types of magazines.*   |  |  |  |  | | --- | --- | --- | --- | | ***home decorating*** | ***news*** | ***health*** | ***computer*** | | *172* | *87* | *82* | *104* | | *286* | *94* | *153* | *136* | | *163* | *123* | *87* | *98* | | *205* | *106* | *103* | *207* | | *197* | *101* | *96* | *146* |   *Table 13.23*  *Using a significance level of 5%, test the hypothesis that the four magazine types have the same mean length.* | |
| Solution | a. *H0*: *μd* = *μn* = *μh* = *μc*  b. Alternative Hypothesis: At least one pair of means is different  c. *df*(*num*) = 3; *df*(*denom*) = 16  d. *F* distribution  e. 8.69  f. 0.0012  g. Check student’s solution.  h. Decision: Reject null hypothesis; Conclusion: There is sufficient evidence to conclude that the mean lengths are different. | |
| Exercise 68. | *Table 13.23**lists the number of pages in four different types of magazines.*   |  |  |  |  | | --- | --- | --- | --- | | ***home decorating*** | ***news*** | ***health*** | ***computer*** | | *172* | *87* | *82* | *104* | | *286* | *94* | *153* | *136* | | *163* | *123* | *87* | *98* | | *205* | *106* | *103* | *207* | | *197* | *101* | *96* | *146* |   *Table 13.23*  *Eliminate one magazine type that you now feel has a mean length different from the others. Redo the hypothesis test, testing that the remaining three means are statistically the same. Use a new solution sheet. Based on this test, are the mean lengths for the remaining three magazines statistically the same?* | |
| Solution | a. *Ha*: *μ*d = *μ*n = *μ*h  b. At least any two of the magazines have different mean lengths.  c. *df*(*num*) = 2, *df*(*denom*) = 12  d. *F* distribtuion  e. *F* = 15.28  f. *p*-value = 0.001  g. Check student’s solution.  h.  i. Alpha: 0.05  ii. Decision: Reject the Null Hypothesis.  iii. Reason for decision: *p*-value < alpha  iv. Conclusion: There is sufficient evidence to conclude that the mean lengths of the magazines are different. | |
| Exercise 69. | *A researcher wants to know if the mean times (in minutes) that people watch their favorite news station are the same. Suppose that Table 13.24**shows the results of a study.*   |  |  |  | | --- | --- | --- | | ***CNN*** | ***FOX*** | ***Local*** | | *45* | *15* | *72* | | *12* | *43* | *37* | | *18* | *68* | *56* | | *38* | *50* | *60* | | *23* | *31* | *51* | | *35* | *22* |  |   *Table 13.24*  *Assume that all distributions are normal, the three population standard deviations are approximately the same, and the data were collected independently and randomly. Use a level of significance of 0.05.* | |
| Solution | a. *H0*: *μc = μf= μl*  b. At least two of the means are different.  c. *df*(*num*) = 2; *df*(*denom*) = 14  d. *F*2,14  e. 4.08  f. 0.0401  g. i. Alpha: 0.05  ii. Decision Reject the null hypothesis  iii. Reason for decision: *p*-value < alpha  iv. Conclusion: There is sufficient evidence to conclude that the mean times are different. | |
| Exercise 70. | *Are the means for the final exams the same for all statistics class delivery types? Table 13.25**shows the scores on final exams from several randomly selected classes that used the different delivery types.*   |  |  |  | | --- | --- | --- | | ***Online*** | ***Hybrid*** | ***Face-to-Face*** | | *72* | *83* | *80* | | *84* | *73* | *78* | | *77* | *84* | *84* | | *80* | *81* | *81* | | *81* |  | *86* | |  |  | *79* | |  |  | *82* |   *Table 13.25*  *Assume that all distributions are normal, the three population standard deviations are approximately the same, and the data were collected independently and randomly. Use a level of significance of 0.05.* | |
| Solution | a. *H0*: μo = μh = μf  b. At least two of the means are different.  c. *df*(*n*) = 2, *df*(d) = 13  d. *F*2,13  e. 0.64  f. 0.5437  g. Check student’s solution.  h:  i. Alpha: 0.05  ii. Decision: Do not reject the null hypothesis.  iii. Reason for decision: *p*-value > alpha  iv: Conclusion: The mean scores of different class delivery are not different. | |
| Exercise 71. | *Are the mean number of times a month a person eats out the same for Whites, Blacks, Hispanics and Asians? Suppose that Table 13.26 shows the results of a study.*   |  |  |  |  | | --- | --- | --- | --- | | ***White*** | ***Black*** | ***Hispanic*** | ***Asian*** | | *6* | *4* | *7* | *8* | | *8* | *1* | *3* | *3* | | *2* | *5* | *5* | *5* | | *4* | *2* | *4* | *1* | | *6* |  | *6* | *7* |   *Table 13.26*  *Assume that all distributions are normal, the four population standard deviations are approximately the same, and the data were collected independently and randomly. Use a level of significance of 0.05.* | |
| Solution | a. *H0* : *μ*w = *μ*b = *μ*a  b. At least two of the means are different.  c. *df*(*num*) = 3; *df*(*denom*) = 15  d. *F*3,15  e. 0.8853  f. 0.4711  g.  i. Alpha: 0.05  ii. Decision: Do not reject the null hypothesis  iii. Reason for decision: *p*-value > alpha  iv. Conclusion: There is insufficient evidence to conclude that the mean numbers of times per month a person eats out are different. | |
| Exercise 72. | *Are the mean numbers of daily visitors to a ski resort the same for the three types of snow conditions? Suppose that Table 13.27 shows the results of a study.*   |  |  |  | | --- | --- | --- | | ***Powder*** | ***Machine Made*** | ***Hard Packed*** | | *1,210* | *2,107* | *2,846* | | *1,080* | *1,149* | *1,638* | | *1,537* | *862* | *2,019* | | *941* | *1,870* | *1,178* | |  | *1,528* | *2,233* | |  | *1,382* |  |   *Table 13.27*  *Assume that all distributions are normal, the three population standard deviations are approximately the same, and the data were collected independently and randomly. Use a level of significance of 0.05.* | |
| Solution | a. *H0*: *μ*p = *μ*m = *μ*h  b. At least any two of the means are different.  c. *df(n)* = 2; *df(d)* =12  d. *F*2,12  e 3.13  f. 0.0807  g. Check student’s solution.  h:  i. Alpha: 0.05  ii. Decision: Do not reject the null hypothesis.  iii. Reason for decision: p-value > alpha  iv: Conclusion: There is not sufficient evidence to conclude that the mean numbers of daily visitors are different. | |
| Exercise 73. | *Sanjay made identical paper airplanes out of three different weights of paper, light, medium and heavy. He made four airplanes from each of the weights, and launched them himself across the room. Here are the distances (in meters) that his planes flew.*   |  |  |  |  |  | | --- | --- | --- | --- | --- | | ***Paper Type/Trial*** | ***Trial 1*** | ***Trial 2*** | ***Trial 3*** | ***Trial 4*** | | *Heavy* | *5.1 meters* | *3.1 meters* | *4.7 meters* | *5.3 meters* | | *Medium* | *4 meters* | *3.5 meters* | *4.5 meters* | *6.1 meters* | | *Light* | *3.1 meters* | *3.3 meters* | *2.1 meters* | *1.9 meters* |   *Table 13.28*  *D:\My Desktop\stat\New Folder\a.jpg*  *Figure 13.8*  *a. Take a look at the data in the graph. Look at the spread of data for each group (light, medium, heavy). Does it seem reasonable to assume a normal distribution with the same variance for each group? Yes or No.*  *b. Why is this a balanced design?*  *c. Calculate the sample mean and sample standard deviation for each group.*  *d. Does the weight of the paper have an effect on how far the plane will travel? Use a 1% level of significance. Complete the test using the method shown in the bean plant example in Example 13.4.*  *◦ variance of the group means \_\_\_\_\_\_\_\_\_\_*  *◦ MSbetween= \_\_\_\_\_\_\_\_\_\_\_*  *◦ mean of the three sample variances \_\_\_\_\_\_\_\_\_\_\_*  *◦ MSwithin = \_\_\_\_\_\_\_\_\_\_\_\_\_*  *◦ F statistic = \_\_\_\_\_\_\_\_\_\_\_\_*  *◦ df(num) = \_\_\_\_\_\_\_\_\_\_, df(denom) = \_\_\_\_\_\_\_\_\_\_\_*  *◦ number of groups \_\_\_\_\_\_\_*  *◦ number of observations \_\_\_\_\_\_\_*  *◦ p-value = \_\_\_\_\_\_\_\_\_\_ (P(F > \_\_\_\_\_\_\_) = \_\_\_\_\_\_\_\_\_\_)*  *◦ Graph the p-value.*  *◦ decision: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_*  *◦ conclusion: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_* | |
| Solution | a. Yes. While the spreads are not identical, they are close enough. Real data are never perfect.  b. Each group has the same number of entries.  C.   |  |  |  |  | | --- | --- | --- | --- | |  | **Light** | **Medium** | **Heavy** | | Sample Mean | 2.6 | 4.525 | 4.55 | | Sample Variance | 0.4933 | 1.2692 | 0.9967 |   d. The variance of the three group means is 1.25146. So *MSbetween* = (4)(1.25146) = 5.00584.  e. The mean of the three sample variances is 0.91973 which is MS*within*.  Finally, the *F* statistic is the ratio of the these two values,  and *df(denom)* = 12 - 3 = 9 where three and 12 are the number of groups and observations respectively. The *p*-value of the test is then *p*-value = *P*(*F* > 5.4427) = 0.0282.    **Decision:** Since the *p*-value is less than the level of significance of 1%, we reject the null hypothesis.  **Conclusion:** We have good evidence from the data that the weight of the paper used to construct the paper airplanes had an effect on the distance the plane traveled. | |
| Exercise 74. | *DDT is a pesticide that has been banned from use in the United States and most other areas of the world. It is quite effective, but persisted in the environment and over time became seen as harmful to higher-level organisms. Famously, egg shells of eagles and other raptors were believed to be thinner and prone to breakage in the nest because of ingestion of DDT in the food chain of the birds.*  *An experiment was conducted on the number of eggs (fecundity) laid by female fruit flies. There are three groups of flies. One group was bred to be resistant to DDT (the RS group). Another was bred to be especially susceptible to DDT (SS). Finally there was a control line of non-selected or typical fruitflies (NS). Here are the data:*   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | ***RS*** | ***SS*** | ***NS*** | ***RS*** | ***SS*** | ***NS*** | | *12.8* | *38.4* | *35.4* | *22.4* | *23.1* | *22.6* | | *21.6* | *32.9* | *27.4* | *27.5* | *29.4* | *40.4* | | *14.8* | *48.5* | *19.3* | *20.3* | *16* | *34.4* | | *23.1* | *20.9* | *41.8* | *38.7* | *20.1* | *30.4* | | *34.6* | *11.6* | *20.3* | *26.4* | *23.3* | *14.9* | | *19.7* | *22.3* | *37.6* | *23.7* | *22.9* | *51.8* | | *22.6* | *30.2* | *36.9* | *26.1* | *22.5* | *33.8* | | *29.6* | *33.4* | *37.3* | *29.5* | *15.1* | *37.9* | | *16.4* | *26.7* | *28.2* | *38.6* | *31* | *29.5* | | *20.3* | *39* | *23.4* | *44.4* | *16.9* | *42.4* | | *29.3* | *12.8* | *33.7* | *23.2* | *16.1* | *36.6* | | *14.9* | *14.6* | *29.2* | *23.6* | *10.8* | *47.4* | | *27.3* | *12.2* | *41.7* |  |  |  |   *Table 13.29*  *The values are the average number of eggs laid daily for each of 75 flies (25 in each group) over the first 14 days of their lives. Using a 1% level of significance, are the mean rates of egg selection for the three strains of fruitfly different? If so, in what way? Specifically, the researchers were interested in whether or not the selectively bred strains were different from the nonselected line, and whether the two selected lines were different from each other. Here is a chart of the three groups:*    Figure 13.9 | |
| Solution | The data appear normally distributed from the chart and of similar spread. There do not appear to be any serious outliers, so we may proceed with our ANOVA calculations, to see if we have good evidence of a difference between the three groups. H0: *μ*1 = *μ*2 = *μ*3; *H*a: *μ*i ≠ *μ*j some *i* ≠ *j*. Define *μ*1, *μ*2, *μ*3, as the population mean number of eggs laid by the three groups of fruit flies. *F* statistic = 8.6657; *p*-value = 0.0004    Figure 13.10  **Decision:** Since the *p*-value is less than the level of significance of 0.01, we reject the null hypothesis.  **Conclusion:** We have good evidence that the average number of eggs laid during the first 14 days of life for these three strains of fruitflies are different. Interestingly, if you perform a two sample *t*-test to compare the RS and NS groups they are significantly different (*p* = 0.0013). Similarly, SS and NS are significantly different (*p* = 0.0006). However, the two selected groups, RS and SS are *not* significantly different (*p* = 0.5176). Thus we appear to have good evidence that selection either for resistance or for susceptibility involves a reduced rate of egg production (for these specific strains) as compared to flies that were not selected for resistance or susceptibility to DDT. Here, genetic selection has apparently involved a loss of fecundity. | |
| Exercise 75. | *The data shown is the recorded body temperatures of 130 subjects as estimated from available histograms. Traditionally we are taught that the normal human body temperature is 98.6°F. This is not quite correct for everyone. Are the mean temperatures among the four groups different?*  *Calculate 95% confidence intervals for the mean body temperature in each group and comment about the confidence intervals.*   |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | | ***FL*** | ***FH*** | ***ML*** | ***MH*** | ***FL*** | ***FH*** | ***ML*** | ***MH*** | | *96.4* | *96.8* | *96.3* | *96.9* | *96.9* | *98.6* | *98.1* | *98.6* | | *96.7* | *97.7* | *96.7* | *97* | *98.7* | *98.6* | *98.1* | *98.6* | | *97.2* | *97.8* | *97.1* | *97.1* | *98.7* | *98.6* | *98.2* | *98.7* | | *97.2* | *97.9* | *97.2* | *97.1* | *98.7* | *98.7* | *98.2* | *98.8* | | *97.4* | *98* | *97.3* | *97.4* | *98.7* | *98.7* | *98.2* | *98.8* | | *97.6* | *98* | *97.4* | *97.5* | *98.8* | *98.8* | *98.2* | *98.8* | | *97.7* | *98* | *97.4* | *97.6* | *98.8* | *98.8* | *98.3* | *98.9* | | *97.8* | *98* | *97.4* | *97.7* | *98.8* | *98.8* | *98.4* | *98.4* | | *97.8* | *98.1* | *97.5* | *97.8* | *98.8* | *98.9* | *98.9* | *99* | | *97.9* | *98.3* | *97.6* | *97.9* | *99.2* | *99* | *98.5* | *99* | | *97.9* | *98.3* | *97.6* | *98* | *99.3* | *99* | *98.5* | *99.2* | | *98* | *98.3* | *97.8* | *98* |  | *99.1* | *98.6* | *99.5* | | *98.2* | *98.4* | *97.8* | *98* |  | *98* | *98* |  | | *98.2* | *98.4* | *97.8* | *98.3* |  | *99.2* | *98.7* |  | | *98.2* | *98.4* | *97.9* | *98.4* |  | *99.4* | *99.1* |  | | *98.2* | *98.4* | *98* | *98.4* |  | *99.9* | *99.3* |  | | *98.2* | *98.5* | *98* | *98.6* |  | *100* | *99.4* |  | | *98.2* | *98.6* | *98* | *98.6* |  | *100.8* |  |  |   Table 13.30 | |
| Solution | The chart of the distribution of body temperatures among the four groups shows no problems. Enter the four groups into four lists in your calculator. Using the ANOVA command in the TESTS section of STAT you should be able to fill in the ANOVA table:   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Source of Variation** | **Sum of Squares (*SS*)** | **Degrees of Freedom (*df*)** | **Mean Square (*MS*)** | ***F*** | | Factor (Between) | 7.452 | 3 | 2.4840 | 5.0569 | | Error (Within) | 61.893 | 126 | 0.4912 |  | | Total | 79130.25 | 129 |  |  |   And the the *p*-value is *P*(*F*3,126 > 5.0569) = 0.0024. **Decision:** The *p*-value is lower than any typical significance level. Reject the null hypothesis of equal means in favor of the alternative hypothesis of at least one mean being different. **Conclusion:** We have very strong evidence of a difference in mean body temperatures among the four groups. Additional Information If we look at 95% confidence intervals for the mean body temperature in each group (Use STAT TESTS 8:TInterval four times, once for each list, on the calculator), we obtain:   |  |  | | --- | --- | | **Group** | **95% Confidence Interval** | | FL | (97.9, 98.4) | | FH | (98.4, 98.8) | | ML | (97.8, 98.2) | | MH | (98.0, 98.5) |   Note that only the women with higher heartbeat rates had a confidence interval for their mean body temperature that includes the traditional 98.6 value. It seems clear, if we believe these data, that the typical person is much more likely than not to have a normal body temperature somewhat below 98.6° F. | |
| Exercise 76. | *Three students, Linda, Tuan, and Javier, are given five laboratory rats each for a nutritional experiment. Each rat’s weight is recorded in grams. Linda feeds her rats Formula A, Tuan feeds his rats Formula B, and Javier feeds his rats Formula C. At the end of a specified time period, each rat is weighted again and the net gain in grams is recorded. Using a significance level of 10%, test the hypothesis that the three formulas produce the same mean weight gain.*   |  |  |  | | --- | --- | --- | | *Linda’s rats* | *Tuan’s rats* | *Javier’s rats* | | *43.5* | *47.0* | *51.2* | | *39.4* | *40.5* | *40.9* | | *41.3* | *38.9* | *37.9* | | *46.0* | *46.3* | *45.0* | | *38.2* | *44.2* | *48.6* |   *Table 13.31*  *Determine whether or not the variance in weight gain is statistically the same among Javier’s and Linda’s rats. Test at a significance level of 10%.* | |
| Solution | a.  b.  c. *df*(*num*) =4; *df*(*denom*) =4  d. *F*4, 4  e. 3.00  f. 2(0.1563) = 0.3126. Using the TI-83+/84+ function 2-SampFtest, you get the the test statistic as 2.9986 and *p*-value directly as 0.3127. If you input the lists in a different order, you get a test statistic of 0.3335 but the *p*-value is the same because this is a two-tailed test.  g. Check student's solution.  h. Decision: Do not reject the null hypothesis; Conclusion: There is insufficient evidence to conclude that the variances are different. | |
| Exercise 77. | *A grassroots group opposed to a proposed increase in the gas tax claimed that the increase would hurt working-class people the most, since they commute the farthest to work. Suppose that the group randomly surveyed 24 individuals and asked them their daily one-way commuting mileage. The results are as follows.*   |  |  |  | | --- | --- | --- | | ***Working-class*** | ***Professional (middle incomes)*** | ***Professional (wealthy)*** | | *17.8* | *16.5* | *8.5* | | *26.7* | *17.4* | *6.3* | | *49.4* | *22.0* | *4.6* | | *9.4* | *7.4* | *12.6* | | *65.4* | *9.4* | *11.0* | | *47.1* | *2.1* | *28.6* | | *19.5* | *6.4* | *15.4* | | *51.2* | *13.9* | *9.3* |   *Table 13.32*  *Determine whether or not the variance in mileage driven is statistically the same among the working-class and professional (middle income) groups. Use a 5% significance level.* | |
| Solution | a.  b.  c. *df(n)* = 7; *df(d)* =7  d. *F*7, 7  e. 9.0  f. 0.010  g. Check student’s solution.  h.  i. Alpha: 0.05  ii. Decision: Reject the null hypothesis.  iii. Reason for decision: *p*-value < alpha  iv. Conclusion: There is sufficient evidence to conclude that the variances are different. | |
| Exercise 78. | *Refer to the data from**Terri Vogel’s Log Book****.*** *Examine practice laps 3 and 4. Determine whether or not the variance in lap time is statistically the same for those practice laps.* | |
| Solution | a.  b.  c. *df(n)* = 19, *df(d)* = 19  d. *F*19, 19  e. 1.13  f. 0.786  g. Check student’s solution.  h.  i. Alpha:0.05  ii. Decision: Do not reject the null hypothesis.  iii. Reason for decision: *p*-value > alpha  iv. Conclusion: There is not sufficient evidence to conclude that the variances are different. | |
| Exercise 79. | *The following table lists the number of pages in four different types of magazines.*   |  |  |  |  | | --- | --- | --- | --- | | ***Home decorating*** | ***News*** | ***Health*** | ***Computer*** | | *172* | *87* | *82* | *104* | | *286* | *94* | *153* | *136* | | *163* | *123* | *87* | *98* | | *205* | *106* | *103* | *207* | | *197* | *101* | *96* | *146* |   *Table 13.33*  *Which two magazine types do you think have the same variance in length?* | |
| Solution | The answers may vary. Sample answer: Home decorating magazines and computer magazines have the same variance. | |
| Exercise 80. | *The following table lists the number of pages in four different types of magazines.*   |  |  |  |  | | --- | --- | --- | --- | | ***Home decorating*** | ***News*** | ***Health*** | ***Computer*** | | *172* | *87* | *82* | *104* | | *286* | *94* | *153* | *136* | | *163* | *123* | *87* | *98* | | *205* | *106* | *103* | *207* | | *197* | *101* | *96* | *146* |   *Table 13.33*  *Which two magazine types do you think have different variances in length?* | |
| Solution | The answers may vary. Sample answer: Home decorating magazines and news magazines have different variances. | |
| Exercise 81. | *Is the variance for the amount of money, in dollars, that shoppers spend on Saturdays at the mall the same as the variance for the amount of money that shoppers spend on Sundays at the mall? Suppose that the Table 13.34**shows the results of a study.*   |  |  |  |  | | --- | --- | --- | --- | | ***Saturday*** | ***Sunday*** | ***Saturday*** | ***Sunday*** | | *75* | *44* | *62* | *137* | | *18* | *58* | *0* | *82* | | *150* | *61* | *124* | *39* | | *94* | *19* | *50* | *127* | | *62* | *99* | *31* | *141* | | *73* | *60* | *118* | *73* | |  | *89* |  |  |   *Table 13.34* | |
| Solution | a.  b.  c. *df(num)* =11*, df(denom)* =12  d. *F*11,12  e. 1.35  f. 0.6090  g. Check student’s solution.  h.  i. Alpha:0.05  ii. Decision: Do not reject the null hypothesis.  iii. Reason for decision: *p*-value > alpha  iv. There is insufficient evidence to conclude that the variances are different. | |
| Exercise 82. | *Are the variances for incomes on the East Coast and the West Coast the same? Suppose that Table 13.35 shows the results of a study. Income is shown in thousands of dollars.*   |  |  | | --- | --- | | ***East*** | ***West*** | | *38* | *71* | | *47* | *126* | | *30* | *42* | | *82* | *51* | | *75* | *44* | | *52* | *90* | | *115* | *88* | | *67* |  |   *Table 13.35*  *Assume that both distributions are normal. Use a level of significance of 0.05.* | |
| Solution | a.  b.  c. *df(n)* = 7, *df(d)* =6  d. *F*7,6  e. 0.8117  f. 0.7825  g. Check student’s solution.  h.  i. Alpha: 0.05  ii. Decision: Do not reject the null hypothesis.  iii. Reason for decision: *p*-value > alpha  iv: Conclusion: There is not sufficient evidence to conclude that the variances are different. | |
| Exercise 83. | *Thirty men in college were taught a method of finger tapping. They were randomly assigned to three groups of ten, with each receiving one of three doses of caffeine: 0 mg, 100 mg, 200 mg. This is approximately the amount in no, one, or two cups of coffee. Two hours after ingesting the caffeine, the men had the rate of finger tapping per minute recorded. The experiment was double blind, so neither the recorders nor the students knew which group they were in. Does caffeine affect the rate of tapping, and if so how?*  *Here are the data:*   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **0mg** | **100mg** | **200mg** | **0mg** | **100mg** | **200mg** | | 242 | 248 | 246 | 245 | 246 | 248 | | 244 | 245 | 250 | 248 | 247 | 252 | | 247 | 248 | 248 | 248 | 250 | 250 | | 242 | 247 | 246 | 244 | 246 | 248 | | 246 | 243 | 245 | 242 | 244 | 250 |   Table 13.36 | |
| Solution | Here is a strip chart of the finger tap data. It is clear that the spread of the three groups is reasonably similar, with no real skew or outliers.    Here are the means and standard deviations for the three groups:   |  |  |  |  | | --- | --- | --- | --- | |  | **0 mg** | **100 mg** | **200 mg** | | Mean | 244.8 | 246.4 | 248.3 | | Standard Deviation | 2.3944 | 2.0656 | 2.2136 | | *n* | 10 | 10 | 10 |   The means go up by about two taps per minute with increasing caffeine dose. Are these differences significant, or could they reasonably happen due to change alone? We construct the ANOVA table:   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Source of Variation** | **Sum of Squares (*SS*)** | **Degrees of Freedom (*df*)** | **Mean Squares (*MS*)** | ***F*** | | Factor (between) | 61.4 | 3 - 1 = 2 |  |  | | Error (within) | 134.1 | 30 - 3 = 27 |  |  | | Total | 195.5 | 30 – 1 = 29 |  |  |   *P*(*x* > 6.1812 ) = 0.0062  At any reasonable and usual level of alpha the test is significant. Reject the null hypothesis. There is sufficient evidence to conclude that the finger taps per minute for college men increase with the caffeine dose. | |
| Exercise 84. | *King Manuel I, Komnenus ruled the Byzantine Empire from Constantinople (Istanbul) during the years 1145 to 1180 A.D. The empire was very powerful during his reign, but declined significantly afterwards. Coins minted during his era were found in Cyprus, an island in the eastern Mediterranean Sea. Nine coins were from his first coinage, seven from the second, four from the third, and seven from a fourth. These spanned most of his reign. We have data on the silver content of the coins:*   |  |  |  |  | | --- | --- | --- | --- | | ***First Coinage*** | ***Second Coinage*** | ***Third Coinage*** | ***Fourth Coinage*** | | *5.9* | *6.9* | *4.9* | *5.3* | | *6.8* | *9.0* | *5.5* | *5.6* | | *6.4* | *6.6* | *4.6* | *5.5* | | *7.0* | *8.1* | *4.5* | *5.1* | | *6.6* | *9.3* |  | *6.2* | | *7.7* | *9.2* |  | *5.8* | | *7.2* | *8.6* |  | *5.8* | | *6.9* |  |  |  | | *6.2* |  |  |  |   *Table 13.37*  *Did the silver content of the coins change over the course of Manuel’s reign?*  *Here are the means and variances of each coinage. The data are unbalanced.*   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | ***First*** | ***Second*** | ***Third*** | ***Fourth*** | | *Mean* | *6.7444* | *8.2429* | *4.875* | *5.6143* | | *Variance* | *0.2953* | *1.2095* | *0.2025* | *0.1314* |   *Table 13.38* | |
| Solution | Here is a strip chart of the silver content of the coins:    Figure 13.11  While there are differences in spread, it is not unreasonable to use ANOVA techniques. Here is the completed ANOVA table:   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Source of Variation** | **Sum of Squares (*SS*)** | **Degrees of Freedom (*df*)** | **Mean Square (*MS*)** | ***F*** | | Factor (Between) | 37.748 | 4 -1 =3 | 12.5825 | 26.272 | | Error (Within) | 11.015 | 27 – 4 = 23 | 0.4789 |  | | Total | 48.763 | 27 – 1 = 26 |  |  |   *P*(*F*>26.272) = 0 Reject the null hypothesis for any alpha. There is sufficient evidence to conclude that the mean silver content among the four coinages is different. From the strip chart, it appears the first and second coinages had higher silver contents than the third and fourth. | |
| Exercise 85. | The American League and the National League of Major League Baseball are each divided into three divisions: East, Central, and West. Many years, fans talk about some divisions being stronger (having better teams) than other divisions. This may have consequences for the postseason. For instance, in 2012 Tampa Bay won 90 games and did not play in the postseason, while Detroit won only 88 and did play in the postseason. This may have been an oddity, but is there good evidence that in the 2012 season, the American League divisions were significantly different in overall records? Use the following data to test whether the mean number of wins per team in the three American League divisions were the same or not. Note that the data are not balanced, as two divisions had five teams, while one had only four.   |  |  |  | | --- | --- | --- | | Division | Team | Wins | | East | NY Yankees | 95 | | East | Baltimore | 93 | | East | Tampa Bay | 90 | | East | Toronto | 73 | | East | Boston | 69 |   **Table 13.39**   |  |  |  | | --- | --- | --- | | **Division** | **Team** | **Wins** | | **Central** | **Detroit** | **88** | | **Central** | **Chicago Sox** | **85** | | **Central** | **Kansas City** | **72** | | **Central** | **Cleveland** | **68** | | **Central** | **Minnesota** | **66** |   **Table 13.40**   |  |  |  | | --- | --- | --- | | **Division** | **Team** | **Wins** | | **West** | **Oakland** | **94** | | **West** | **Texas** | **93** | | **West** | **LA Angels** | **89** | | **West** | **Seattle** | **75** |   **Table 13.41** | |
| Solution | Here is a strip chart of the number of wins for the 14 teams in the AL for the 2012 season.    Figure 13.12  While the spread seems similar, there may be some question about the normality of the data, given the wide gaps in the middle near the 0.500 mark of 82 games (teams play 162 games each season in MLB). However, One-Way ANOVA is robust.  Here is the ANOVA table for the data:   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Source of**  **Variation** | **Sum of Squares (*SS*)** | **Degrees of**  **Freedom (*df*)** | **Mean Square**  **(*MS*)** | ***F*** | | Factor (Between) |  |  | 172.08 | 1.5521 | | Error (Within) |  |  | 110.87 |  | | Total | 1563.71 |  |  |  |   *Table 13.43*  *P*(*F* > 1.5521) = 0.2548  Since the p-value is so large, there is not good evidence against the null hypothesis of equal means. We fail to reject the null hypothesis. Thus, for 2012, there is not any good evidence of a significant difference in mean number of wins between the divisions of the American League. | |

This file is copyright 2015, Rice University. All Rights Reserved.