Chapter 3

**Experiments with a Single Factor: The Analysis of Variance**

# Solutions

**3.1.** An experimenter has conducted a single-factor experiment with four levels of the factor, and each factor level has been replicated six times. The computed value of the *F*-statistic is *F0* = 3.26. Find bounds on the *P*-value.

Table *P*-value = 0.025, 0.050 Computer *P*-value = 0.043

**3.2.** An experimenter has conducted a single-factor experiment with six levels of the factor, and each factor level has been replicated three times. The computed value of the *F*-statistic is *F0* = 5.81. Find bounds on the *P*-value.

Table *P*-value < 0.010 Computer *P*-value = 0.006

**3.3S.** An experiment has conducted a single-factor completely randomized design with five levels of the factor and three replicates. The computed value of the *F*-statistic is 4.87. Find bounds on the *P*-value.

*N*= 5 factors levels x 3 replicates = 15

Degrees of freedom for the factor: *a* – 1 = 5 – 1 = 4

Degrees of freedom Total = 15 – 1= 14

Degrees of freedom error = Total – factor = 14 – 4 = 10

Bounds of *P*-value for *F* = 4.87 with 4 and 10 degrees of freedom are = 0.025 < *P* < 0.01

**3.4.** An experiment has conducted a single-factor completely randomized design with three levels of factors and five replicates. The computed value of the *F*-statistic is 2.91. Find bounds on the *P*-value.

*N*= 3 factors levels x 5 replicates = 15

Degrees of freedom for the factor: *a* – 1 = 3 – 1 = 2

Degrees of freedom Total = 15 – 1= 14

Degrees of freedom error = Total – factor = 14 – 2 = 12

Bounds of *P*-value for *F* = 2.91 with 2 and 12 degrees of freedom are = 0.01 < *P* < 0.05

**3.5.** The mean square for error in the ANOVA provides an estimate of

**(a) The variance of the random error**

(b) The variance of an individual treatment average

(c) The standard deviation of an individual observation

(d) None of the above

**3.6S.** It is always a good idea to check the normality assumption in the ANOVA by applying a test for normality such as the Anderson-Darling test to the residuals.

True **False**

**3.7.** A computer ANOVA output is shown below. Fill in the blanks. You may give bounds on the *P*-value.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| One-way ANOVA | | | | | |
| Source | DF | SS | MS | F | P |
| Factor | ? | ? | 246.93 | ? | ? |
| Error | 25 | 186.53 | ? |  |  |
| Total | 29 | 1174.24 |  |  |  |

Completed table is:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| One-way ANOVA | | | | | |
| Source | DF | SS | MS | F | P |
| Factor | 4 | 987.71 | 246.93 | 33.09 | < 0.0001 |
| Error | 25 | 186.53 | 7.46 |  |  |
| Total | 29 | 1174.24 |  |  |  |

**3.8.** An article appeared in *The Wall Street Journal* on Tuesday, April 27, 2010, with the title “Eating Chocolate Is Linked to Depression.” The article reported on a study funded by the National Heart, Lung and Blood Institute (part of the National Institutes of Health) and conducted by the faculty at the University of California, San Diego, and the University of California, Davis. The research was also published in the *Archives of Internal Medicine* (2010, pp. 699-703). The study examined 931 adults who were not taking antidepressants and did not have known cardiovascular disease or diabetes. The group was about 70% men and the average age of the group was reported to be about 58. The participants were asked about chocolate consumption and then screened for depression using a questionnaire. People who scored less than 16 on the questionnaire are not considered depressed, while those with scores above 16 and less than or equal to 22 are considered possibly depressed, while those with scores above 22 are considered likely to be depressed. The survey found that people who were not depressed ate an average of 8.4 servings of chocolate per month, while those individuals who scored above 22 were likely to be depressed ate the most chocolate, an average of 11.8 servings per month. No differentiation was made between dark and milk chocolate. Other foods were also examined, but no patterned emerged between other foods and depression. Is this study really a designed experiment? Does it establish a cause-and-effect link between chocolate consumption and depression? How would the study have to be conducted to establish such a link?

This is not a designed experiment, and it does not establish a cause-and-effect link between chocolate consumption and depression. An experiment could be run by giving a group of people a defined amount of chocolate servings per month for several months, while not giving another group any chocolate. Ideally it would be good to have the participants not eat any chocolate for a period of time before the experiment, and measure depression for each participant before and after the experiment.

**3.9.** The tensile strength of Portland cement is being studied. Four different mixing techniques can be used economically. A completely randomized experiment was conducted and the following data were collected.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mixing Technique | Tensile Strength (lb/in2) | | | |
| 1 | 3129 | 3000 | 2865 | 2890 |
| 2 | 3200 | 3300 | 2975 | 3150 |
| 3 | 2800 | 2900 | 2985 | 3050 |
| 4 | 2600 | 2700 | 2600 | 2765 |

(a) Test the hypothesis that mixing techniques affect the strength of the cement. Use *α* = 0.05.

Design Expert Output

**Response:** **Tensile Strength** **in lb/in^2**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 4.897E+005 3 1.632E+005 12.73 0.0005 significant

*A* *4.897E+005* *3* *1.632E+005* *12.73* *0.0005*

Residual 1.539E+005 12 12825.69

*Lack of Fit* *0.000* *0*

*Pure Error* *1.539E+005* *12* *12825.69*

Cor Total 6.436E+005 15

The Model F-value of 12.73 implies the model is significant. There is only

a 0.05% chance that a "Model F-Value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-1 2971.00 56.63

2-2 3156.25 56.63

3-3 2933.75 56.63

4-4 2666.25 56.63

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 -185.25 1 80.08 -2.31 0.0392

1 vs 3 37.25 1 80.08 0.47 0.6501

1 vs 4 304.75 1 80.08 3.81 0.0025

2 vs 3 222.50 1 80.08 2.78 0.0167

2 vs 4 490.00 1 80.08 6.12 < 0.0001

3 vs 4 267.50 1 80.08 3.34 0.0059

The *F*-value is 12.73 with a corresponding *P*-value of .0005. Mixing technique has an effect.

(b) Construct a graphical display as described in Section 3.5.3 to compare the mean tensile strengths for the four mixing techniques. What are your conclusions?





Based on examination of the plot, we would conclude that  and  are the same; that  differs from  and , that  differs from  and , and that  and  are different.

(c) Use the Fisher LSD method with **=0.05 to make comparisons between pairs of means.



Treatment 2 vs. Treatment 4 = 3156.250 - 2666.250 = 490.000 > 174.495

Treatment 2 vs. Treatment 3 = 3156.250 - 2933.750 = 222.500 > 174.495

Treatment 2 vs. Treatment 1 = 3156.250 - 2971.000 = 185.250 > 174.495

Treatment 1 vs. Treatment 4 = 2971.000 - 2666.250 = 304.750 > 174.495

Treatment 1 vs. Treatment 3 = 2971.000 - 2933.750 = 37.250 < 174.495

Treatment 3 vs. Treatment 4 = 2933.750 - 2666.250 = 267.500 > 174.495

The Fisher LSD method is also presented in the Design-Expert computer output above. The results agree with the graphical method for this experiment.

(d) Construct a normal probability plot of the residuals. What conclusion would you draw about the validity of the normality assumption?

There is nothing unusual about the normal probability plot of residuals.



(e) Plot the residuals versus the predicted tensile strength. Comment on the plot.

There is nothing unusual about this plot.



(f) Prepare a scatter plot of the results to aid the interpretation of the results of this experiment.

Design-Expert automatically generates the scatter plot. The plot below also shows the sample average for each treatment and the 95 percent confidence interval on the treatment mean.



**3.10.**

(a) Rework part (c) of Problem 3.9 using Tukey’s test with ** = 0.05. Do you get the same conclusions from Tukey’s test that you did from the graphical procedure and/or the Fisher LSD method?

Minitab Output

Tukey's pairwise comparisons

Family error rate = 0.0500

Individual error rate = 0.0117

Critical value = 4.20

Intervals for (column level mean) - (row level mean)

1 2 3

2 -423

53

3 -201 -15

275 460

4 67 252 30

543 728 505

No, the conclusions are not the same. The mean of Treatment 4 is different than the means of Treatments 1, 2, and 3. However, the mean of Treatment 2 is not different from the means of Treatments 1 and 3 according to Tukey’s method, they were found to be different using the graphical method and the Fisher LSD method.

(b) Explain the difference between the Tukey and Fisher procedures.

Both Tukey and Fisher utilize a single critical value; however, Tukey’s is based on the studentized range statistic while Fisher’s is based on *t* distribution.

**3.11S.** Reconsider the experiment in Problem 3.9. Find a 95 percent confidence interval on the mean tensile strength of the Portland cement produced by each of the four mixing techniques. Also find a 95 percent confidence interval on the difference in means for techniques 1 and 3. Does this aid in interpreting the results of the experiment?



Treatment 1: 





Treatment 2: 3156.25±123.387



Treatment 3: 2933.75±123.387



Treatment 4: 2666.25±123.387



Treatment 1 - Treatment 3: 





Because the confidence interval for the difference between means 1 and 3 spans zero, we agree with the statement in Problem 3.9 (b); there is not a statistical difference between these two means.

**3.12.** A product developer is investigating the tensile strength of a new synthetic fiber that will be used to make cloth for men’s shirts. Strength is usually affected by the percentage of cotton used in the blend of materials for the fiber. The engineer conducts a completely randomized experiment with five levels of cotton content and replicated the experiment five times. The data are shown in the following table.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Cotton  Weight Percentage | Observations | | | | |
| 15 | 7 | 7 | 15 | 11 | 9 |
| 20 | 12 | 17 | 12 | 18 | 18 |
| 25 | 14 | 19 | 19 | 18 | 18 |
| 30 | 19 | 25 | 22 | 19 | 23 |
| 35 | 7 | 10 | 11 | 15 | 11 |

(a) Is there evidence to support the claim that cotton content affects the mean tensile strength? Use *α* = 0.05.

Minitab Output

**One-way ANOVA: Tensile Strength versus Cotton Percentage**

Analysis of Variance for Tensile

Source DF SS MS F P

Cotton P 4 475.76 118.94 14.76 0.000

Error 20 161.20 8.06

Total 24 636.96

Yes, the *F*-value is 14.76 with a corresponding *P*-value of 0.000. The percentage of cotton in the fiber appears to have an affect on the tensile strength.

(b) Use the Fisher LSD method to make comparisons between the pairs of means. What conclusions can you draw?

Minitab Output

Fisher's pairwise comparisons

Family error rate = 0.264

Individual error rate = 0.0500

Critical value = 2.086

Intervals for (column level mean) - (row level mean)

15 20 25 30

20 -9.346

-1.854

25 -11.546 -5.946

-4.054 1.546

30 -15.546 -9.946 -7.746

-8.054 -2.454 -0.254

35 -4.746 0.854 3.054 7.054

2.746 8.346 10.546 14.546

In the Minitab output the pairs of treatments that do not contain zero in the pair of numbers indicates that there is a difference in the pairs of the treatments. 15% cotton is different than 20%, 25% and 30%. 20% cotton is different than 30% and 35% cotton. 25% cotton is different than 30% and 35% cotton. 30% cotton is different than 35%.

(c) Analyze the residuals from this experiment and comment on model adequacy.

The residual plots below show nothing unusual.





**3.13.** Reconsider the experiment described in Problem 3.12. Suppose that 30 percent cotton content is a control. Use Dunnett’s test with *α* = 0.05 to compare all of the other means with the control.

For this problem: *a* = 5, *a*-1 = 4, *f*=20, *n*=5 and *α* = 0.05





The control treatment, treatment 4, differs from treatments 1, 2 and 5.

**3.14S.** A pharmaceutical manufacturer wants to investigate the bioactivity of a new drug. A completely randomized single-factor experiment was conducted with three dosage levels, and the following results were obtained.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dosage | Observations | | | |
| 20g | 24 | 28 | 37 | 30 |
| 30g | 37 | 44 | 31 | 35 |
| 40g | 42 | 47 | 52 | 38 |

(a) Is there evidence to indicate that dosage level affects bioactivity? Use *α* = 0.05.

Minitab Output

**One-way ANOVA: Activity versus Dosage**

Analysis of Variance for Activity

Source DF SS MS F P

Dosage 2 450.7 225.3 7.04 0.014

Error 9 288.3 32.0

Total 11 738.9

There appears to be a different in the dosages.

(b) If it is appropriate to do so, make comparisons between the pairs of means. What conclusions can you draw?

Because there appears to be a difference in the dosages, the comparison of means is appropriate.

Minitab Output

Tukey's pairwise comparisons

Family error rate = 0.0500

Individual error rate = 0.0209

Critical value = 3.95

Intervals for (column level mean) - (row level mean)

20g 30g

30g -18.177

4.177

40g -26.177 -19.177

-3.823 3.177

The Tukey comparison shows a difference in the means between the 20g and the 40g dosages.

(c) Analyze the residuals from this experiment and comment on the model adequacy.

There is nothing too unusual about the residual plots shown below.





**3.15.** A rental car company wants to investigate whether the type of car rented affects the length of the rental period. An experiment is run for one week at a particular location, and 10 rental contracts are selected at random for each car type. The results are shown in the following table.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Type of Car | Observations | | | | | | | | | |
| Sub-compact | 3 | 5 | 3 | 7 | 6 | 5 | 3 | 2 | 1 | 6 |
| Compact | 1 | 3 | 4 | 7 | 5 | 6 | 3 | 2 | 1 | 7 |
| Midsize | 4 | 1 | 3 | 5 | 7 | 1 | 2 | 4 | 2 | 7 |
| Full Size | 3 | 5 | 7 | 5 | 10 | 3 | 4 | 7 | 2 | 7 |

(a) Is there evidence to support a claim that the type of car rented affects the length of the rental contract? Use *α* = 0.05. If so, which types of cars are responsible for the difference?

Minitab Output

**One-way ANOVA: Days versus Car Type**

Analysis of Variance for Days

Source DF SS MS F P

Car Type 3 16.68 5.56 1.11 0.358

Error 36 180.30 5.01

Total 39 196.98

There is no difference.

(b) Analyze the residuals from this experiment and comment on the model adequacy.





There is nothing unusual about the residuals.

(c) Notice that the response variable in this experiment is a count. Should this cause any potential concerns about the validity of the analysis of variance?

Because the data is count data, a square root transformation could be applied. The analysis is shown below. It does not change the interpretation of the data.

Minitab Output

**One-way ANOVA: Sqrt Days versus Car Type**

Analysis of Variance for Sqrt Day

Source DF SS MS F P

Car Type 3 1.087 0.362 1.10 0.360

Error 36 11.807 0.328

Total 39 12.893

**3.16S.** I belong to a golf club in my neighborhood. I divide the year into three golf seasons: summer (June-September), winter (November-March) and shoulder (October, April and May). I believe that I play my best golf during the summer (because I have more time and the course isn’t crowded) and shoulder (because the course isn’t crowded) seasons, and my worst golf during the winter (because all of the part-year residents show up, and the course is crowded, play is slow, and I get frustrated). Data from the last year are shown in the following table.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Season | Observations | | | | | | | | | |
| Summer | 83 | 85 | 85 | 87 | 90 | 88 | 88 | 84 | 91 | 90 |
| Shoulder | 91 | 87 | 84 | 87 | 85 | 86 | 83 |  |  |  |
| Winter | 94 | 91 | 87 | 85 | 87 | 91 | 92 | 86 |  |  |

(a) Do the data indicate that my opinion is correct? Use *α* = 0.05.

Minitab Output

**One-way ANOVA: Score versus Season**

Analysis of Variance for Score

Source DF SS MS F P

Season 2 35.61 17.80 2.12 0.144

Error 22 184.63 8.39

Total 24 220.24

The data do not support the author’s opinion.

(b) Analyze the residuals from this experiment and comment on model adequacy.





There is nothing unusual about the residuals.

**3.17.** A regional opera company has tried three approaches to solicit donations from 24 potential sponsors. The 24 potential sponsors were randomly divided into three groups of eight, and one approach was used for each group. The dollar amounts of the resulting contributions are shown in the following table.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Approach | Contributions (in $) | | | | | | | |
| 1 | 1000 | 1500 | 1200 | 1800 | 1600 | 1100 | 1000 | 1250 |
| 2 | 1500 | 1800 | 2000 | 1200 | 2000 | 1700 | 1800 | 1900 |
| 3 | 900 | 1000 | 1200 | 1500 | 1200 | 1550 | 1000 | 1100 |

(a) Do the data indicate that there is a difference in results obtained from the three different approaches? Use *α* = 0.05.

Minitab Output

**One-way ANOVA: Contribution versus Approach**

Analysis of Variance for Contribution

Source DF SS MS F P

Approach 2 1362708 681354 9.41 0.001

Error 21 1520625 72411

Total 23 2883333

There is a difference between the approaches. The Tukey test will indicate which are different. Approach 2 is different than approach 1 and approach 3.

Minitab Output

Tukey's pairwise comparisons

Family error rate = 0.0500

Individual error rate = 0.0200

Critical value = 3.56

Intervals for (column level mean) - (row level mean)

1 2

2 -770

-93

3 -214 218

464 895

(b) Analyze the residuals from this experiment and comment on the model adequacy.





There is nothing unusual about the residuals.

**3.18.** An experiment was run to determine whether four specific firing temperatures affect the density of a certain type of brick. A completely randomized experiment led to the following data:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Temperature | Density | | | | |
| 100 | 21.8 | 21.9 | 21.7 | 21.6 | 21.7 |
| 125 | 21.7 | 21.4 | 21.5 | 21.4 |  |
| 150 | 21.9 | 21.8 | 21.8 | 21.6 | 21.5 |
| 175 | 21.9 | 21.7 | 21.8 | 21.4 |  |

(a) Does the firing temperature affect the density of the bricks? Use *α* = 0.05.

No, firing temperature does not affect the density of the bricks. Refer to the Design-Expert output below.

Design Expert Output

**Response:** **Density**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 0.16 3 0.052 2.02 0.1569 not significant

*A* *0.16* *3* *0.052* *2.02* *0.1569*

Residual 0.36 14 0.026

*Lack of Fit* *0.000* *0*

*Pure Error* *0.36* *14* *0.026*

Cor Total 0.52 17

The "Model F-value" of 2.02 implies the model is not significant relative to the noise. There is a

15.69 % chance that a "Model F-value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-100 21.74 0.072

2-125 21.50 0.080

3-150 21.72 0.072

4-175 21.70 0.080

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 0.24 1 0.11 2.23 0.0425

1 vs 3 0.020 1 0.10 0.20 0.8465

1 vs 4 0.040 1 0.11 0.37 0.7156

2 vs 3 -0.22 1 0.11 -2.05 0.0601

2 vs 4 -0.20 1 0.11 -1.76 0.0996

3 vs 4 0.020 1 0.11 0.19 0.8552

(b) Is it appropriate to compare the means using the Fisher LSD method in this experiment?

The analysis of variance tells us that there is no difference in the treatments. There is no need to proceed with Fisher’s LSD method to decide which mean is difference.

(c) Analyze the residuals from this experiment. Are the analysis of variance assumptions satisfied? There is nothing unusual about the residual plots.



(d) Construct a graphical display of the treatments as described in Section 3.5.3. Does this graph adequately summarize the results of the analysis of variance in part (b). Yes.



**3.19.** Rework Part (d) of Problem 3.18 using the Tukey method. What conclusions can you draw? Explain carefully how you modified the procedure to account for unequal sample sizes.

When sample sizes are unequal, the appropriate formula for the Tukey method is



Treatment 1 vs. Treatment 2 = 21.74 – 21.50 = 0.24 < 0.314

Treatment 1 vs. Treatment 3 = 21.74 – 21.72 = 0.02 < 0.296

Treatment 1 vs. Treatment 4 = 21.74 – 21.70 = 0.04 < 0.314

Treatment 3 vs. Treatment 2 = 21.72 – 21.50 = 0.22 < 0.314

Treatment 4 vs. Treatment 2 = 21.70 – 21.50 = 0.20 < 0.331

Treatment 3 vs. Treatment 4 = 21.72 – 21.70 = 0.02 < 0.314

All pairwise comparisons do not identify differences. Notice that there are different critical values for the comparisons depending on the sample sizes of the two groups being compared.

Because we could not reject the hypothesis of equal means using the analysis of variance, we should never have performed the Tukey test (or any other multiple comparison procedure, for that matter). If you ignore the analysis of variance results and run multiple comparisons, you will likely make type I errors.

**3.20.** A manufacturer of television sets is interested in the effect of tube conductivity of four different types of coating for color picture tubes. A completely randomized experiment is conducted and the following conductivity data are obtained:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Coating Type | Conductivity | | | |
| 1 | 143 | 141 | 150 | 146 |
| 2 | 152 | 149 | 137 | 143 |
| 3 | 134 | 136 | 132 | 127 |
| 4 | 129 | 127 | 132 | 129 |

(a) Is there a difference in conductivity due to coating type? Use *α* = 0.05.

Yes, there is a difference in means. Refer to the Design-Expert output below..

Design Expert Output

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 844.69 3 281.56 14.30 0.0003 significant

*A* *844.69* *3* *281.56* *14.30* *0.0003*

Residual 236.25 12 19.69

*Lack of Fit* *0.000* *0*

*Pure Error* *236.25* *12* *19.69*

Cor Total 1080.94 15

The Model F-value of 14.30 implies the model is significant. There is only

a 0.03% chance that a "Model F-Value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-1 145.00 2.22

2-2 145.25 2.22

3-3 132.25 2.22

4-4 129.25 2.22

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 -0.25 1 3.14 -0.080 0.9378

1 vs 3 12.75 1 3.14 4.06 0.0016

1 vs 4 15.75 1 3.14 5.02 0.0003

2 vs 3 13.00 1 3.14 4.14 0.0014

2 vs 4 16.00 1 3.14 5.10 0.0003

3 vs 4 3.00 1 3.14 0.96 0.3578

(b) Estimate the overall mean and the treatment effects.



(c) Compute a 95 percent interval estimate of the mean of coating type 4. Compute a 99 percent interval estimate of the mean difference between coating types 1 and 4.

Treatment 4: 



Treatment 1 - Treatment 4: 



(d) Test all pairs of means using the Fisher LSD method with *α*=0.05.

Refer to the Design-Expert output above. The Fisher LSD procedure is automatically included in the output.

The means of Coating Type 2 and Coating Type 1 are not different. The means of Coating Type 3 and Coating Type 4 are not different. However, Coating Types 1 and 2 produce higher mean conductivity than does Coating Types 3 and 4.

(e) Use the graphical method discussed in Section 3.5.3 to compare the means. Which coating produces the highest conductivity?

 Coating types 1 and 2 produce the highest conductivity.



(f) Assuming that coating type 4 is currently in use, what are your recommendations to the manufacturer? We wish to minimize conductivity.

Since coatings 3 and 4 do not differ, and as they both produce the lowest mean values of conductivity, use either coating 3 or 4. As type 4 is currently being used, there is probably no need to change.

**3.21S.** Reconsider the experiment in Problem 3.20. Analyze the residuals and draw conclusions about model adequacy.

There is nothing unusual in the normal probability plot. A funnel shape is seen in the plot of residuals versus predicted conductivity indicating a possible non-constant variance.





**3.22S.** An article in *Environment International* (Vol. 18, No. 4, 1992) describes an experiment in which the amount of radon released in showers was investigated. Radon ­enriched water was used in the experiment and six different orifice diameters were tested in shower heads. The data from the experiment are shown in the following table.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Orifice  Diameter | Radon Released (%) | | | |
| 0.37 | 80 | 83 | 83 | 85 |
| 0.51 | 75 | 75 | 79 | 79 |
| 0.71 | 74 | 73 | 76 | 77 |
| 1.02 | 67 | 72 | 74 | 74 |
| 1.40 | 62 | 62 | 67 | 69 |
| 1.99 | 60 | 61 | 64 | 66 |

(a) Does the size of the orifice affect the mean percentage of radon released? Use *α* = 0.05.

Yes. There is at least one treatment mean that is different.

Design Expert Output

**Response:** **Radon Released** **in %**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 1133.38 5 226.68 30.85 < 0.0001 significant

*A* *1133.38* *5* *226.68* *30.85* *< 0.0001*

Residual 132.25 18 7.35

*Lack of Fit* *0.000* *0*

*Pure Error* *132.25* *18* *7.35*

Cor Total 1265.63 23

The Model F-value of 30.85 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-0.37 82.75 1.36

2-0.51 77.00 1.36

3-0.71 75.00 1.36

4-1.02 71.75 1.36

5-1.40 65.00 1.36

6-1.99 62.75 1.36

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 5.75 1 1.92 3.00 0.0077

1 vs 3 7.75 1 1.92 4.04 0.0008

1 vs 4 11.00 1 1.92 5.74 < 0.0001

1 vs 5 17.75 1 1.92 9.26 < 0.0001

1 vs 6 20.00 1 1.92 10.43 < 0.0001

2 vs 3 2.00 1 1.92 1.04 0.3105

2 vs 4 5.25 1 1.92 2.74 0.0135

2 vs 5 12.00 1 1.92 6.26 < 0.0001

2 vs 6 14.25 1 1.92 7.43 < 0.0001

3 vs 4 3.25 1 1.92 1.70 0.1072

3 vs 5 10.00 1 1.92 5.22 < 0.0001

3 vs 6 12.25 1 1.92 6.39 < 0.0001

4 vs 5 6.75 1 1.92 3.52 0.0024

4 vs 6 9.00 1 1.92 4.70 0.0002

5 vs 6 2.25 1 1.92 1.17 0.2557

(b) Find the P-value for the F statistic in part (a).

*P*=3.161 x 10-8

(c) Analyze the residuals from this experiment.

There is nothing unusual about the residuals.





(d) Find a 95 percent confidence interval on the mean percent radon released when the orifice diameter is 1.40.

Treatment 5 (Orifice =1.40): 



(e) Construct a graphical display to compare the treatment means as describe in Section 3­.5.3. What conclusions can you draw?



Treatments 5 and 6 as a group differ from the other means; 2, 3, and 4 as a group differ from the other means, 1 differs from the others.

**3.23.** The response time in milliseconds was determined for three different types of circuits that could be used in an automatic valve shutoff mechanism. The results are shown in the following table.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Circuit Type | Response Time | | | | |
| 1 | 9 | 12 | 10 | 8 | 15 |
| 2 | 20 | 21 | 23 | 17 | 30 |
| 3 | 6 | 5 | 8 | 16 | 7 |

(a) Test the hypothesis that the three circuit types have the same response time. Use *α* = 0.01.

From the computer printout, *F*=16.08, so there is at least one circuit type that is different.

Design Expert Output

**Response:** **Response Time** **in ms**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 543.60 2 271.80 16.08 0.0004 significant

*A* *543.60* *2* *271.80* *16.08* *0.0004*

Residual 202.80 12 16.90

*Lack of Fit* *0.000* *0*

*Pure Error* *202.80* *12* *16.90*

Cor Total 746.40 14

The Model F-value of 16.08 implies the model is significant. There is only

a 0.04% chance that a "Model F-Value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-1 10.80 1.84

2-2 22.20 1.84

3-3 8.40 1.84

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 -11.40 1 2.60 -4.38 0.0009

1 vs 3 2.40 1 2.60 0.92 0.3742

2 vs 3 13.80 1 2.60 5.31 0.0002

(b) Use Tukey’s test to compare pairs of treatment means. Use *α* = 0.01.







1 vs. 2: ⏐10.8-22.2⏐=11.4 > 9.266

1 vs. 3: ⏐10.8-8.4⏐=2.4 < 9.266

2 vs. 3: ⏐22.2-8.4⏐=13.8 > 9.266

1 and 2 are different. 2 and 3 are different.

Notice that the results indicate that the mean of treatment 2 differs from the means of both treatments 1 and 3, and that the means for treatments 1 and 3 are the same. Notice also that the Fisher LSD procedure (see the computer output) gives the same results.

(c) Use the graphical procedure in Section 3.5.3 to compare the treatment means. What conclusions can you draw? How do they compare with the conclusions from part (a).

The scaled-*t* plot agrees with part (b). In this case, the large difference between the mean of treatment 2 and the other two treatments is very obvious.



(d) Construct a set of orthogonal contrasts, assuming that at the outset of the experiment you suspected the response time of circuit type 2 to be different from the other two.





Type 2 differs from the average of type 1 and type 3.

(e) If you were a design engineer and you wished to minimize the response time, which circuit type would you select?

Either type 1 or type 3 as they are not different from each other and have the lowest response time.

(f) Analyze the residuals from this experiment. Are the basic analysis of variance assumptions satisfied?

The normal probability plot has some points that do not lie along the line in the upper region. This may indicate potential outliers in the data.





**3.24S.** The effective life of insulating fluids at an accelerated load of 35 kV is being studied. Test data have been obtained for four types of fluids. The results from a completely randomized experiment were as follows:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Fluid Type | | Life (in h) at 35 kV Load | | | | | | |
| 1 | 17.6 | | 18.9 | 16.3 | 17.4 | 20.1 | 21.6 |
| 2 | 16.9 | | 15.3 | 18.6 | 17.1 | 19.5 | 20.3 |
| 3 | 21.4 | | 23.6 | 19.4 | 18.5 | 20.5 | 22.3 |
| 4 | 19.3 | | 21.1 | 16.9 | 17.5 | 18.3 | 19.8 |

(a) Is there any indication that the fluids differ? Use *α* = 0.05.

At *α* = 0.05 there is no difference, but since the *P*-value is just slightly above 0.05, there is probably a difference in means.

Design Expert Output

**Response:** **Life** **in in h**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 30.17 3 10.06 3.05 0.0525 not significant

*A* *30.16* *3* *10.05* *3.05* *0.0525*

Residual 65.99 20 3.30

*Lack of Fit* *0.000* *0*

*Pure Error* *65.99* *20* *3.30*

Cor Total 96.16 23

The Model F-value of 3.05 implies there is a 5.25% chance that a "Model F-Value"

this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-1 18.65 0.74

2-2 17.95 0.74

3-3 20.95 0.74

4-4 18.82 0.74

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 0.70 1 1.05 0.67 0.5121

1 vs 3 -2.30 1 1.05 -2.19 0.0403

1 vs 4 -0.17 1 1.05 -0.16 0.8753

2 vs 3 -3.00 1 1.05 -2.86 0.0097

2 vs 4 -0.87 1 1.05 -0.83 0.4183

3 vs 4 2.13 1 1.05 2.03 0.0554

(b) Which fluid would you select, given that the objective is long life?

Treatment 3. The Fisher LSD procedure in the computer output indicates that the fluid 3 is different from the others, and it’s average life also exceeds the average lives of the other three fluids.

(c) Analyze the residuals from this experiment. Are the basic analysis of variance assumptions satisfied? There is nothing unusual in the residual plots.





**3.25.** Four different designs for a digital computer circuit are being studied in order to compare the amount of noise present. The following data have been obtained:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Circuit Design | Noise Observed | | | | |
| 1 | 19 | 20 | 19 | 30 | 8 |
| 2 | 80 | 61 | 73 | 56 | 80 |
| 3 | 47 | 26 | 25 | 35 | 50 |
| 4 | 95 | 46 | 83 | 78 | 97 |

(a) Is the amount of noise present the same for all four designs? Use α = 0.05.

No, at least one treatment mean is different.

Design Expert Output

**Response:** **Noise**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 12042.00 3 4014.00 21.78 < 0.0001 significant

*A* *12042.00* *3* *4014.00* *21.78* *< 0.0001*

Residual 2948.80 16 184.30

*Lack of Fit* *0.000* *0*

*Pure Error* *2948.80* *16* *184.30*

Cor Total 14990.80 19

The Model F-value of 21.78 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-1 19.20 6.07

2-2 70.00 6.07

3-3 36.60 6.07

4-4 79.80 6.07

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 -50.80 1 8.59 -5.92 < 0.0001

1 vs 3 -17.40 1 8.59 -2.03 0.0597

1 vs 4 -60.60 1 8.59 -7.06 < 0.0001

2 vs 3 33.40 1 8.59 3.89 0.0013

2 vs 4 -9.80 1 8.59 -1.14 0.2705

3 vs 4 -43.20 1 8.59 -5.03 0.0001

(b) Analyze the residuals from this experiment. Are the basic analysis of variance assumptions satisfied? There is nothing too unusual about the residual plots, although there is a mild outlier present.





(c) Which circuit design would you select for use? Low noise is best.

From the Design Expert Output, the Fisher LSD procedure comparing the difference in means identifies Type 1 as having lower noise than Types 2 and 4. Although the LSD procedure comparing Types 1 and 3 has a *P*-value greater than 0.05, it is less than 0.10. Unless there are other reasons for choosing Type 3, Type 1 would be selected.

**3.26.** Four chemists are asked to determine the percentage of methyl alcohol in a certain chemical compound. Each chemist makes three determinations, and the results are the following:

|  |  |  |  |
| --- | --- | --- | --- |
| Chemist | Percentage of Methyl Alcohol | | |
| 1 | 84.99 | 84.04 | 84.38 | |
| 2 | 85.15 | 85.13 | 84.88 | |
| 3 | 84.72 | 84.48 | 85.16 | |
| 4 | 84.20 | 84.10 | 84.55 | |

(a) Do chemists differ significantly? Use *α* = 0.05.

There is no significant difference at the 5% level, but chemists differ significantly at the 10% level.

Design Expert Output

**Response:** **Methyl Alcohol** **in %**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 1.04 3 0.35 3.25 0.0813 not significant

*A* *1.04* *3* *0.35* *3.25* *0.0813*

Residual 0.86 8 0.11

*Lack of Fit* *0.000* *0*

*Pure Error* *0.86* *8* *0.11*

Cor Total 1.90 11

The Model F-value of 3.25 implies there is a 8.13% chance that a "Model F-Value"

this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-1 84.47 0.19

2-2 85.05 0.19

3-3 84.79 0.19

4-4 84.28 0.19

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 -0.58 1 0.27 -2.18 0.0607

1 vs 3 -0.32 1 0.27 -1.18 0.2703

1 vs 4 0.19 1 0.27 0.70 0.5049

2 vs 3 0.27 1 0.27 1.00 0.3479

2 vs 4 0.77 1 0.27 2.88 0.0205

3 vs 4 0.50 1 0.27 1.88 0.0966

(b) Analyze the residuals from this experiment.

There is nothing unusual about the residual plots.





(c) If chemist 2 is a new employee, construct a meaningful set of orthogonal contrasts that might have been useful at the start of the experiment.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Chemists | Total | C1 | C2 | C3 |
| 1 | 253.41 | 1 | -2 | 0 |
| 2 | 255.16 | -3 | 0 | 0 |
| 3 | 254.36 | 1 | 1 | -1 |
| 4 | 252.85 | 1 | 1 | 1 |
|  | Contrast Totals: | -4.86 | 0.39 | -1.51 |



Only contrast 1 is significant at 5%.

**3.27S.** Three brands of batteries are under study. It is s suspected that the lives (in weeks) of the three brands are different. Five randomly selected batteries of each brand are tested with the following results:

|  |  |  |
| --- | --- | --- |
| Weeks of Life | | |
| Brand 1 | Brand 2 | Brand 3 |
| 100 | 76 | 108 |
| 96 | 80 | 100 |
| 92 | 75 | 96 |
| 96 | 84 | 98 |
| 92 | 82 | 100 |

(a) Are the lives of these brands of batteries different?

Yes, at least one of the brands is different.

Design Expert Output

**Response:** **Life** **in Weeks**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 1196.13 2 598.07 38.34 < 0.0001 significant

*A* *1196.13* *2* *598.07* *38.34* *< 0.0001*

Residual 187.20 12 15.60

*Lack of Fit* *0.000* *0*

*Pure Error* *187.20* *12* *15.60*

Cor Total 1383.33 14

The Model F-value of 38.34 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-1 95.20 1.77

2-2 79.40 1.77

3-3 100.40 1.77

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 15.80 1 2.50 6.33 < 0.0001

1 vs 3 -5.20 1 2.50 -2.08 0.0594

2 vs 3 -21.00 1 2.50 -8.41 < 0.0001

(b) Analyze the residuals from this experiment.

There is nothing unusual about the residuals.





(c) Construct a 95 percent interval estimate on the mean life of battery brand 2. Construct a 99 percent interval estimate on the mean difference between the lives of battery brands 2 and 3.



Brand 2: 





Brand 2 - Brand 3: 





(d) Which brand would you select for use? If the manufacturer will replace without charge any battery that fails in less than 85 weeks, what percentage would the company expect to replace?

Chose brand 3 for longest life. Mean life of this brand in 100.4 weeks, and the variance of life is estimated by 15.60 (*MSE*). Assuming normality, then the probability of failure before 85 weeks is:



That is, about 5 out of 100,000 batteries will fail before 85 weeks.

**3.28.** Four catalysts that may affect the concentration of one component in a three­ component liquid mixture are being investigated. The following concentrations are obtained from a completely randomized experiment:

|  |  |  |  |
| --- | --- | --- | --- |
| Catalyst | | | |
| 1 | 2 | 3 | 4 |
| 58.2 | 56.3 | 50.1 | 52.9 |
| 57.2 | 54.5 | 54.2 | 49.9 |
| 58.4 | 57.0 | 55.4 | 50.0 |
| 55.8 | 55.3 |  | 51.7 |
| 54.9 |  |  |  |

(a) Do the four catalysts have the same effect on concentration?

No, their means are different.

Design Expert Output

**Response:** **Concentration**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 85.68 3 28.56 9.92 0.0014 significant

*A* *85.68* *3* *28.56* *9.92* *0.0014*

Residual 34.56 12 2.88

*Lack of Fit* *0.000* *0*

*Pure Error* *34.56* *12* *2.88*

Cor Total 120.24 15

The Model F-value of 9.92 implies the model is significant. There is only

a 0.14% chance that a "Model F-Value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-1 56.90 0.76

2-2 55.77 0.85

3-3 53.23 0.98

4-4 51.13 0.85

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 1.13 1 1.14 0.99 0.3426

1 vs 3 3.67 1 1.24 2.96 0.0120

1 vs 4 5.77 1 1.14 5.07 0.0003

2 vs 3 2.54 1 1.30 1.96 0.0735

2 vs 4 4.65 1 1.20 3.87 0.0022

3 vs 4 2.11 1 1.30 1.63 0.1298

(b) Analyze the residuals from this experiment.

There is nothing unusual about the residual plots.





(c) Construct a 99 percent confidence interval estimate of the mean response for catalyst 1.



Catalyst 1: 





**3.29.** A semiconductor manufacturer has developed three different methods for reducing particle counts on wafers. All three methods are tested on five wafers and the after-treatment particle counts obtained. The data are shown below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Method | Count | | | | | |
| 1 | 31 | 10 | 21 | 4 | 1 |
| 2 | 62 | 40 | 24 | 30 | 35 |
| 3 | 58 | 27 | 120 | 97 | 68 |

(a) Do all methods have the same effect on mean particle count?

No, at least one method has a different effect on mean particle count.

Design Expert Output

**Response:** **Count**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 8963.73 2 4481.87 7.91 0.0064 significant

*A* *8963.73* *2* *4481.87* *7.91* *0.0064*

Residual 6796.00 12 566.33

*Lack of Fit* *0.000* *0*

*Pure Error* *6796.00* *12* *566.33*

Cor Total 15759.73 14

The Model F-value of 7.91 implies the model is significant. There is only

a 0.64% chance that a "Model F-Value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-1 13.40 10.64

2-2 38.20 10.64

3-3 73.00 10.64

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 -24.80 1 15.05 -1.65 0.1253

1 vs 3 -59.60 1 15.05 -3.96 0.0019

2 vs 3 -34.80 1 15.05 -2.31 0.0393

(b) Plot the residuals versus the predicted response. Construct a normal probability plot of the residuals. Are there potential concerns about the validity of the assumptions?

The plot of residuals versus predicted appears to be funnel shaped. This indicates the variance of the original observations is not constant. The residuals plotted in the normal probability plot do not fall along a straight line, which suggests that the normality assumption is not valid. A data transformation is recommended.



(c) Based on your answer to part (b) conduct another analysis of the particle count data and draw appropriate conclusions.

For count data, a square root transformation is often very effective in resolving problems with inequality of variance. The analysis of variance for the transformed response is shown below. The difference between methods is much more apparent after applying the square root transformation.

Design Expert Output

**Response:** **Count** **Transform:** **Square root** **Constant:** **0.000**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 63.90 2 31.95 9.84 0.0030 significant

*A* *63.90* *2* *31.95* *9.84* *0.0030*

Residual 38.96 12 3.25

*Lack of Fit* *0.000* *0*

*Pure Error* *38.96* *12* *3.25*

Cor Total 102.86 14

The Model F-value of 9.84 implies the model is significant. There is only

a 0.30% chance that a "Model F-Value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-1 3.26 0.81

2-2 6.10 0.81

3-3 8.31 0.81

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 -2.84 1 1.14 -2.49 0.0285

1 vs 3 -5.04 1 1.14 -4.42 0.0008

2 vs 3 -2.21 1 1.14 -1.94 0.0767

**3.30S.** Several ovens in a metal working shop are used to heat metal specimens. All ovens are supposed to operate at the same temperature, although it is suspected that this may not be true. Three ovens selected at random, and their temperatures on successive heats are noted. The data collected are as follows:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Oven | Temperature | | | | | |
| 1 | 491.50 | 498.30 | 498.10 | 493.50 | 493.60 |  |
| 2 | 488.50 | 484.65 | 479.90 | 477.35 |  |  |
| 3 | 480.10 | 484.80 | 488.25 | 473.00 | 471.85 | 478.65 |

(a) Is there significant variation in temperature between ovens? Use *α*=0.05.

The computer output below shows that there is oven to oven variation.

Minitab Output

**ANOVA: Temp versus Oven**

Factor Type Levels Values

Oven random 3 1, 2, 3

Analysis of Variance for Temp

Source DF SS MS F P

Oven 2 705.10 352.55 13.33 0.001

Error 12 317.31 26.44

Total 14 1022.41

S = 5.14224 R-Sq = 68.96% R-Sq(adj) = 63.79%

(b) Estimate the components of variation for this model.



(c) Analyze the residuals from this experiment and draw conclusions about model adequacy.

There are no concerns with the residual plots below.



**3.31.** An article in the *Journal of the Electrochemical Society* (Vol. 139, No. 2, 1992, pp. 524-532) describes an experiment to investigate low-pressure vapor deposition of polysilicon. The experiment was carried out in a large capacity reactor at Sematech in Austin, Texas. The reactor has several wafer positions, and four of these positions are selected at random. The response variable is film thickness uniformity. Three replicates of the experiment were run, and the data are as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| Wafer Positions | Uniformity | | |
| 1 | 2.76 | 5.67 | 4.49 |
| 2 | 1.43 | 1.70 | 2.19 |
| 3 | 2.34 | 1.97 | 1.47 |
| 4 | 0.94 | 1.36 | 1.65 |

(a) Is there a difference in the wafer positions? Use Use *α*=0.05.

The JMP output below identifies a difference in the wafer positions.

JMP Output

**Summary of Fit**

|  |  |
| --- | --- |
| RSquare | 0.756617 |
| RSquare Adj | 0.665349 |
| Root Mean Square Error | 0.807579 |
| Mean of Response | 2.330833 |
| Observations (or Sum Wgts) | 12 |

**Analysis of Variance**

| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Ratio** |
| --- | --- | --- | --- | --- |
| Models | 3 | 16.219825 | 5.40661 | 8.2900 |
| Error | 8 | 5.217467 | 0.65218 | Prob > F |
| C. Total | 11 | 21.437292 |  | 0.0077\* |

**Effect Tests**

| **Source** | **Nparm** | **DF** | **Sum of Squares** | **F Ratio** | **Prob > F** |  |
| --- | --- | --- | --- | --- | --- | --- |
| Wafer Position | 3 | 3 | 16.219825 | 8.2900 | 0.0077\* |  |

(b) Estimate the variability due to wafer position.

The JMP REML output below identifies the variance component for the wafer position as 1.5848.

JMP Output

**Parameter Estimates**

| **Term** |  | **Estimate** | **Std Error** | **DFDen** | **t Ratio** | **Prob>|t|** |
| --- | --- | --- | --- | --- | --- | --- |
| Intercept |  | 2.3308333 | 0.671231 | 3 | 3.47 | 0.0403\* |

**REML Variance Component Estimates**

| **Random Effect** | **Var Ratio** | **Var Component** | **Std Error** | **95% Lower** | **95% Upper** | **Pct of Total** |
| --- | --- | --- | --- | --- | --- | --- |
| Wafer Position | 2.4300043 | 1.5848083 | 1.4755016 | -1.307122 | 4.4767383 | 70.846 |
| Residual |  | 0.6521833 | 0.3260917 | 0.2975536 | 2.393629 | 29.154 |
| Total |  | 2.2369917 |  |  |  | 100.000 |

**Covariance Matrix of Variance Component Estimates**

| **Random Effect** | **Wafer Position** | **Residual** |
| --- | --- | --- |
| Wafer Position | 2.177105 | -0.035445 |
| Residual | -0.035445 | 0.1063358 |

(c) Estimate the random error component.

The JMP REML output above identifies the random error variance component as 0.6522..

(d) Analyze the residuals from this experiment and comment on model adequacy.

The plot of residuals vs. predicted shows some uniformity concerns.



The residuals used in the plot below are based on the REML analysis. The normal plot shows some concerns with the normality assumption; however, the normality is not important for this analysis.



Uniformity data often requires a transformation, such as a log transformation, and should be considered for this experiment.

**3.32.** Consider the vapor-deposition experiment described in Problem 3.31.

(a) Estimate the total variability in the uniformity response.

The JMP REML output shown in part (b) of Problem 3.31 identifies the total variability as 2.2370.

(b) How much of the total variability in the uniformity response is due to the difference between positions in the reactor?

From the JMP REML output shown in part (b) of Problem 3.31, the differences between positions represents 70.846% of the total variability.

(c) To what level could the variability in the uniformity response be reduced if position-to-position variability in the reactor could be eliminated? Do you believe this is a significant reduction?

The variability could be reduced to 29.154% of the current total variability. Based on the 95% confidence intervals calculated below, this is not significant. An increase in sample size might reverse this decision.



**3.33.** A single-factor completely randomized design has four levels of the factor. There are three replicates and the total sum of squares is 330.56. The treatment sum of squares is 250.65.

(a) What is the estimate of the error variance *σ*2?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| One-way ANOVA | | | | | |
| Source | DF | SS | MS | F | P |
| Factor | 3 | 250.65 | 83.55 | 8.36 | < 0.01 |
| Error | 8 | 79.91 | 9.99 |  |  |
| Total | 11 | 330.56 |  |  |  |



(b) What proportion of the variability in the response variable is explained by the treatment effect?

By calculating the treatment variance component,



*R*2 is “loosely” interpreted as the proportion of the variability explained by the ANOVA model, so



**3.34.** A single-factor completely randomized design has six levels of the factor. There are five replicates and the total sum of squares is 900.025. The treatment sum of squares is 750.50.

(a) What is the estimate of the error variance *σ*2?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| One-way ANOVA | | | | | |
| Source | DF | SS | MS | F | P |
| Factor | 5 | 750.50 | 150.10 | 24.06 | < 0.0001 |
| Error | 24 | 149.75 | 6.24 |  |  |
| Total | 29 | 900.25 |  |  |  |



(b) What proportion of the variability in the response variable is explained by the treatment effect?

By calculating the treatment variance component,



*R*2 is “loosely” interpreted as the proportion of the variability explained by the ANOVA model, so



**3.35.** Find a 95% confidence interval on the interclass correlation coefficient for the experiment in Problem 3.33.



**3.36.** Find a 95% confidence interval on the interclass correlation coefficient for the experiment in Problem 3.34.



**3.37.** Consider testing the equality of the means of two normal populations, where the variances are unknown but are assumed to be equal. The appropriate test procedure is the pooled *t* test. Show that the pooled *t* test is equivalent to the single factor analysis of variance.

 assuming *n*1 = *n*2 = *n*

** for a=2

Furthermore, , which is exactly the same as SSTreatments in a one-way classification with a=2. Thus we have shown that . In general, we know that  so that . Thus the square of the test statistic from the pooled *t*-test is the same test statistic that results from a single-factor analysis of variance with a=2.

**3.38S.** Show that the variance of the linear combination  is .

, 



**3.39.** In a fixed effects experiment, suppose that there are *n* observations for each of four treatments. Let  be single-degree-of-freedom components for the orthogonal contrasts. Prove that .





 and since

, we have 

for a=4.

**3.40.** Use Bartlett's test to determine if the assumption of equal variances is satisfied in Problem 3.27S. Use ** = 0.05. Did you reach the same conclusion regarding the equality of variance by examining the residual plots?

, where



 





 

Cannot reject null hypothesis; conclude that the variance are equal. This agrees with the residual plots in Problem 3.27S.

**3.41.** Use the modified Levene test to determine if the assumption of equal variances is satisfied on Problem 3.27S. Use ** = 0.05. Did you reach the same conclusion regarding the equality of variances by examining the residual plots?

The absolute value of Battery Life – brand median is:

|  |  |  |
| --- | --- | --- |
|  | | |
| Brand 1 | Brand 2 | Brand 3 |
| 4 | 4 | 8 |
| 0 | 0 | 0 |
| 4 | 5 | 4 |
| 0 | 4 | 2 |
| 4 | 2 | 0 |

The analysis of variance indicates that there is not a difference between the different brands and therefore the assumption of equal variances is satisfied.

Design Expert Output

**Response:** **Mod Levine**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 0.93 2 0.47 0.070 0.9328

*A* *0.93* *2* *0.47* *0.070* *0.9328*

Pure Error 80.00 12 6.67

Cor Total 80.93 14

**3.42.** Refer to Problem 3.23. If we wish to detect a maximum difference in mean response times of 10 milliseconds with a probability of at least 0.90, what sample size should be used? How would you obtain a preliminary estimate of ?

Minitab Output



Choose *n* ≥ 6, therefore *N* ≥ 18

Notice that we have used an estimate of the variance obtained from the present experiment. This indicates that we probably didn’t use a large enough sample (*n* was 5 in problem 3.23) to satisfy the criteria specified in this problem. However, the sample size *was* adequate to detect differences in one of the circuit types.

When we have no prior estimate of variability, sometimes we will generate sample sizes for a range of possible variances to see what effect this has on the size of the experiment. Often a knowledgeable expert will be able to bound the variability in the response, by statements such as “the standard deviation is going to be *at least*…” or “the standard deviation shouldn’t be larger than…”.

**3.43.** Refer to Problem 3.23.

(a) If we wish to detect a maximum difference in mean battery life of 10 hours with a probability of at least 0.90, what sample size should be used? Discuss how you would obtain a preliminary estimate of **2 for answering this question.

Minitab Output



Choose *n* ≥ 6, therefore *N* ≥ 18

See the discussion from the previous problem about the estimate of variance.

(b) If the maximum difference between brands is 8 hours, what sample size should be used if we wish to detect this with a probability of at least 0.90?

Minitab Output



Choose *n* ≥ 8, therefore *N* ≥ 24

**3.44.** Consider the experiment in Problem 3.27. If we wish to construct a 95 percent confidence interval on the difference in two mean battery lives that has an accuracy of ±2 weeks, how many batteries of each brand must be tested?

  



Trial and Error yields:

|  |  |  |  |
| --- | --- | --- | --- |
| N |  | *t* | width |
| 5 | 12 | 2.179 | 5.44 |
| 10 | 27 | 2.05 | 3.62 |
| 31 | 90 | 1.99 | 1.996 |
| 32 | 93 | 1.99 | 1.96 |

Choose *n* ≥ 31, therefore *N* ≥ 93

**3.45.** Suppose that four normal populations have means of *μ*1=50, *μ*2=60, *μ*3=50, and *μ*4=60. How many observations should be taken from each population so that the probability of rejecting the null hypothesis of equal population means is at least 0.90? Assume that *α*=0.05 and that a reasonable estimate of the error variance is  =25.

Minitab Output



Therefore, select *n* ≥ 9

**3.46.** Refer to Problem 3.45.

(a) How would your answer change if a reasonable estimate of the experimental error variance were  = 36?

Minitab Output

Therefore, select *n* ≥ 11

(b) How would your answer change if a reasonable estimate of the experimental error variance were  = 49?

Minitab Output

 Therefore, select *n* ≥ 15

(c) Can you draw any conclusions about the sensitivity of your answer in the particular situation about how your estimate of *σ* affects the decision about sample size?

As our estimate of variability increases the sample size must increase to ensure the same power of the test.

(d) Can you make any recommendations about how we should use this general approach to choosing *n* in practice?

When we have no prior estimate of variability, sometimes we will generate sample sizes for a range of possible variances to see what effect this has on the size of the experiment. Often a knowledgeable expert will be able to bound the variability in the response, by statements such as “the standard deviation is going to be *at least*…” or “the standard deviation shouldn’t be larger than…”.

**3.47.** Refer to the aluminum smelting experiment described in Section 3.8.3. Verify that ratio control methods do not affect average cell voltage. Construct a normal probability plot of residuals. Plot the residuals versus the predicted values. Is there an indication that any underlying assumptions are violated?

Design Expert Output

**Response:** **Cell Average**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 2.746E-003 3 9.153E-004 0.20 0.8922 not significant

*A* *2.746E-003* *3* *9.153E-004* *0.20* *0.8922*

Residual 0.090 20 4.481E-003

*Lack of Fit* *0.000* *0*

*Pure Error* *0.090* *20* *4.481E-003*

Cor Total 0.092 23

The "Model F-value" of 0.20 implies the model is not significant relative to the noise. There is a

89.22 % chance that a "Model F-value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-1 4.86 0.027

2-2 4.83 0.027

3-3 4.85 0.027

4-4 4.84 0.027

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 0.027 1 0.039 0.69 0.4981

1 vs 3 0.013 1 0.039 0.35 0.7337

1 vs 4 0.025 1 0.039 0.65 0.5251

2 vs 3 -0.013 1 0.039 -0.35 0.7337

2 vs 4 -1.667E-003 1 0.039 -0.043 0.9660

3 vs 4 0.012 1 0.039 0.30 0.7659

The following residual plots are satisfactory.





**3.48S.** Refer to the aluminum smelting experiment in Section 3.8.3. Verify the ANOVA for pot noise summarized in Table 3.17. Examine the usual residual plots and comment on the experimental validity.

Design Expert Output

**Response:** **Cell StDev** **Transform:** **Natural log** **Constant:** **0.000**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 6.17 3 2.06 21.96 < 0.0001 significant

*A* *6.17* *3* *2.06* *21.96* *< 0.0001*

Residual 1.87 20 0.094

*Lack of Fit* *0.000* *0*

*Pure Error* *1.87* *20* *0.094*

Cor Total 8.04 23

The Model F-value of 21.96 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-1 -3.09 0.12

2-2 -3.51 0.12

3-3 -2.20 0.12

4-4 -3.36 0.12

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 0.42 1 0.18 2.38 0.0272

1 vs 3 -0.89 1 0.18 -5.03 < 0.0001

1 vs 4 0.27 1 0.18 1.52 0.1445

2 vs 3 -1.31 1 0.18 -7.41 < 0.0001

2 vs 4 -0.15 1 0.18 -0.86 0.3975

3 vs 4 1.16 1 0.18 6.55 < 0.0001

The following residual plots identify the residuals to be normally distributed, randomly distributed through the range of prediction, and uniformly distributed across the different algorithms. This validates the assumptions for the experiment.





**3.49.** Four different feed rates were investigated in an experiment on a CNC machine producing a component part used in an aircraft auxiliary power unit. The manufacturing engineer in charge of the experiment knows that a critical part dimension of interest may be affected by the feed rate. However, prior experience has indicated that only dispersion effects are likely to be present. That is, changing the feed rate does not affect the average dimension, but it could affect dimensional variability. The engineer makes five production runs at each feed rate and obtains the standard deviation of the critical dimension (in 10-3 mm). The data are shown below. Assume that all runs were made in random order.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Feed Rate |  | Production | Run |  |  |
| (in/min) | 1 | 2 | 3 | 4 | 5 |
| 10 | 0.09 | 0.10 | 0.13 | 0.08 | 0.07 |
| 12 | 0.06 | 0.09 | 0.12 | 0.07 | 0.12 |
| 14 | 0.11 | 0.08 | 0.08 | 0.05 | 0.06 |
| 16 | 0.19 | 0.13 | 0.15 | 0.20 | 0.11 |

(a) Does feed rate have any effect on the standard deviation of this critical dimension?

Because the residual plots were not acceptable for the non-transformed data, a square root transformation was applied to the standard deviations of the critical dimension. Based on the computer output below, the feed rate has an effect on the standard deviation of the critical dimension.

Design Expert Output

**Response:** **Run StDev** **Transform:** **Square root** **Constant:** **0.000**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 0.040 3 0.013 7.05 0.0031 significant

*A* *0.040* *3* *0.013* *7.05* *0.0031*

Residual 0.030 16 1.903E-003

*Lack of Fit* *0.000* *0*

*Pure Error* *0.030* *16* *1.903E-003*

Cor Total 0.071 19

The Model F-value of 7.05 implies the model is significant. There is only

a 0.31% chance that a "Model F-Value" this large could occur due to noise.

**Treatment Means (Adjusted, If Necessary)**

**Estimated** **Standard**

**Mean** **Error**

1-10 0.30 0.020

2-12 0.30 0.020

3-14 0.27 0.020

4-16 0.39 0.020

**Mean** **Standard** **t for H0**

**Treatment** **Difference** **DF** **Error** **Coeff=0** **Prob > |t|**

1 vs 2 4.371E-003 1 0.028 0.16 0.8761

1 vs 3 0.032 1 0.028 1.15 0.2680

1 vs 4 -0.088 1 0.028 -3.18 0.0058

2 vs 3 0.027 1 0.028 0.99 0.3373

2 vs 4 -0.092 1 0.028 -3.34 0.0042

3 vs 4 -0.12 1 0.028 -4.33 0.0005

(b) Use the residuals from this experiment of investigate model adequacy. Are there any problems with experimental validity?

The residual plots are satisfactory.





**3.50.** Consider the data shown in Problem 3.23.

(a) Write out the least squares normal equations for this problem, and solve them for  and , using the usual constraint . Estimate .



Imposing , therefore , , , 



(b) Solve the equations in (a) using the constraint . Are the estimators  and  the same as you found in (a)? Why? Now estimate  and compare your answer with that for (a). What statement can you make about estimating contrasts in the ?

Imposing the constraint,  we get the following solution to the normal equations: , , , and . These estimators are not the same as in part (a). However, , is the same as in part (a). The contrasts are estimable.

(c) Estimate ,  and  using the two solutions to the normal equations. Compare the results obtained in each case.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Contrast | Estimated from Part (a) | Estimated from Part (b) |
| 1 |  | 10.80 | 10.80 |
| 2 |  | -9.00 | -9.00 |
| 3 |  | 19.20 | 24.60 |

Contrasts 1 and 2 are estimable, 3 is not estimable.

**3.51.** Apply the general regression significance test to the experiment in Example 3.6. Show that the procedure yields the same results as the usual analysis of variance.

From the etch rate table:



from Example 3.6, we have:



, with 20 degrees of freedom.



with 4 degrees of freedom.



with 20-4 degrees of freedom.

This is identical to the SSE found in Example 3.6.

The reduced model:

, with 1 degree of freedom.

, with 4-1=3 degrees of freedom.

Note:  from Example 3.1.

Finally,



which is the same as computed in Example 3.6.

**3.52S.** Use the Kruskal-Wallis test for the experiment in Problem 3.24. Are the results comparable to those found by the usual analysis of variance?

From Design Expert Output of Problem 3.21

**Response:** **Life** **in in h**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 30.17 3 10.06 3.05 0.0525 not significant

*A* *30.16* *3* *10.05* *3.05* *0.0525*

Residual 65.99 20 3.30

*Lack of Fit* *0.000* *0*

*Pure Error* *65.99* *20* *3.30*

Cor Total 96.16 23





Accept the null hypothesis; the treatments are not different. This agrees with the analysis of variance.

**3.52.** Use the Kruskal-Wallis test for the experiment in Problem 3.25. Compare conclusions obtained with those from the usual analysis of variance?

From Design Expert Output of Problem 3.22

**Response:** **Noise**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 12042.00 3 4014.00 21.78 < 0.0001 significant

*A* *12042.00* *3* *4014.00* *21.78* *< 0.0001*

Residual 2948.80 16 184.30

*Lack of Fit* *0.000* *0*

*Pure Error* *2948.80* *16* *184.30*

Cor Total 14990.80 19





Reject the null hypothesis because the treatments are different. This agrees with the analysis of variance.

**3.53.** Consider the experiment in Example 3.6. Suppose that the largest observation on etch rate is incorrectly recorded as 250A/min. What effect does this have on the usual analysis of variance? What effect does it have on the Kruskal-Wallis test?

The incorrect observation reduces the analysis of variance *F*0 from 66.8 to 0.50. It does change the value of the Kruskal-Wallis test statistic but not the result.

Minitab Output

**One-way ANOVA: Etch Rate 2 versus Power**

Analysis of Variance for Etch Rat

Source DF SS MS F P

Power 3 15927 5309 0.50 0.685

Error 16 168739 10546

Total 19 184666

**3.54.** The normality assumption is extremely important in the analysis of variance.

True **False**

**3.55.** The analysis of variance treats both quantitative and qualitative factors alike so far as the basic computations for sums of squares are concerned.

**True** False

**3.56.** If a single-factor experiment has *a* levels of factor and a polynomial of degree *a* – 1 is fit to the experimental data, the error sum of squares for the polynomial model with be exactly the same as the error sums of square for the standard ANOVA.

**True** False

**3.57.** Fisher’s LSD procedure is an extremely conservative method for comparing pairs of treatment means following an ANOVA.

True **False**

**3.58.** The REML method of estimating variance components is a technique based on maximum likelihood, while the ANOVA method is a method-of-moments procedure.

**True** False

**3.59.** One advantage of the REML method of estimating variance components is that it automatically produces confidence intervals on the variance components.

**True** False

**3.60.** The Tukey method is used to compare all treatment means to a control.

True **False**

**3.61.** The estimate of the standard deviation of any observation in the experiment in problem 3.70 is

(a) 7.03

**(b) 2.65**

(c) 5.91

(d) 1.95

(e) none of the above