STAT 8120 – Module 3 Homework

Due 1/26/2020

Connor Armstrong

***3.11*** *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:*

**3.11 Conditions**

|  |
| --- |
| *3.112 Omit parts (b) and (f) and Levene’s for part (e)*  *3.113 For the rest of the semester, always give residual analysis for any problem in SG3 Table 7 order. Use RA (Residual Analysis) comparisons to justify the need for a transformation (with vs. without). Report multiple Tukey comparisons using a PC Bar Plot, discuss sigma distances. Give factor levels producing an optimum.*  *3.11+ Use both SAS and Minitab. Do not repeat discussions, circle one common numeric value in outputs.* |

**Table 3.11.1**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 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 |

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

The data in Table 3.11.1 was processed using Minitab. The following ANOVA table was included in the results:

**Figure 3.11.a.1 Minitab ANOVA Table**

**Analysis of Variance**

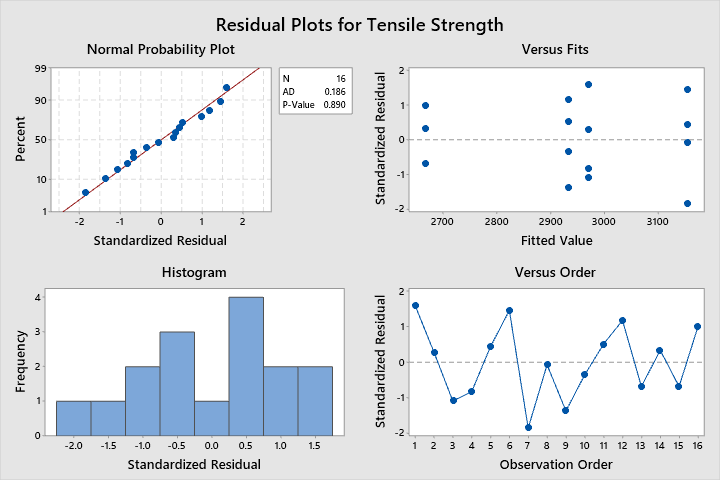
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Adj SS** | **Adj MS** | **F-Value** | **P-Value** |
| Mixing Technique | 3 | 489740 | 163247 | 12.73 | 0.000 |
| Error | 12 | 153908 | 12826 |  |  |
| Total | 15 | 643648 |  |  |  |

Given a P-value of 0.000 from the ANOVA table above, and given a significance level of 0.05, the null hypothesis ought to be rejected. Therefore, it is unlikely that the means for the 4 groups in table 3.11.1 are equivalent. It can be said with 95% certainty that at least one of the means is not equivalent to the others.

It is important to check the assumptions for any statistical analysis. Page 16 of Study Guide 3 for this course contains a “Summary of Model Adequacy Checking” which will be utilized as an outline to ensure all assumptions are addressed properly (This satisfies condition 3.113).

**Assumption 1, Normality** – The p-value for the A.D. test for normality for this data is 0.890 (Not Significant), as seen near the Normal probability plot in the Minitab output below. Given a significance level of 0.05, this p-value will result in a failure to reject the null hypothesis that the data is normally distributed. In other words, the data is normally distributed, with a confidence level of 95%.

**Figure 3.11.a.2 Minitab Residual Analysis**



**Assumption 2, Constant Variance** – Levene’s test would verify this assumption, but the condition 3.112 states “*Omit…Levene’s for part (e)*.” Therefore, the assumption verification will be omitted.

**Assumption 3, No Outliers** – The data was checked for outliers using Minitab. There are no outliers in this data. See Minitab output below:

**Figure 3.11.a.3 Minitab Residual Analysis**

**Outlier Test: Tensile Strength**

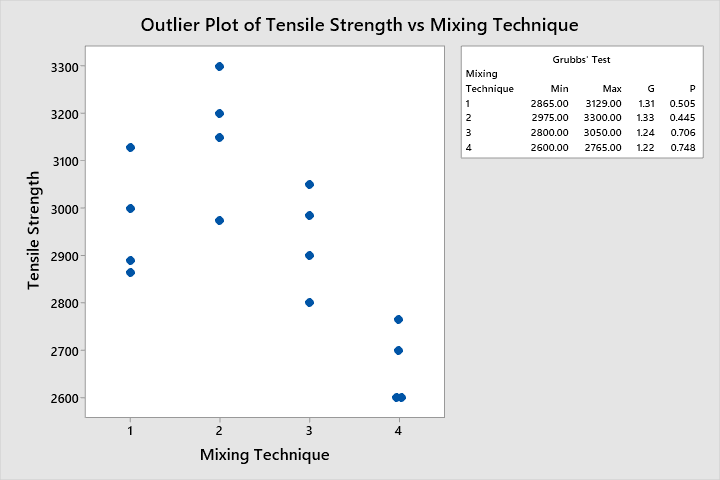
**Method**

|  |  |
| --- | --- |
| Null hypothesis | All data values come from the same normal population |
| Alternative hypothesis | Smallest or largest data value is an outlier |
| Significance level | α = 0.05 |

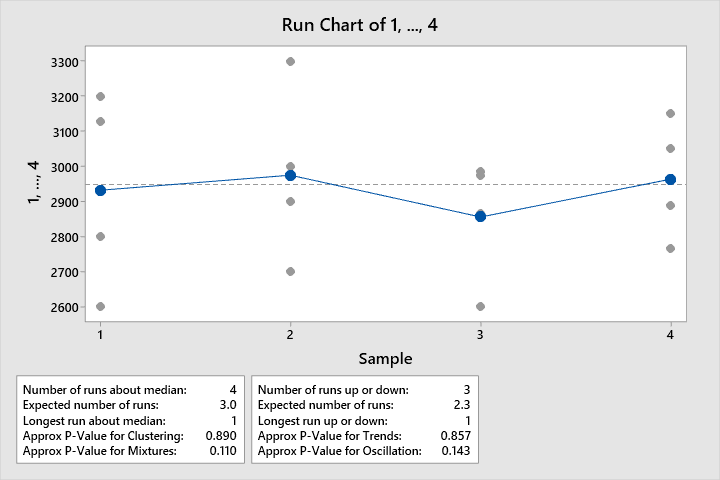
**Grubbs' Test**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Mixing Technique** | **N** | **Mean** | **StDev** | **Min** | **Max** | **G** | **P** |
| Tensile Strength | 1 | 4 | 2971.0 | 120.6 | 2865.0 | 3129.0 | 1.31 | 0.505 |
|  | 2 | 4 | 3156.3 | 136.0 | 2975.0 | 3300.0 | 1.33 | 0.445 |
|  | 3 | 4 | 2933.8 | 108.3 | 2800.0 | 3050.0 | 1.24 | 0.706 |
|  | 4 | 4 | 2666.3 | 81.0 | 2600.0 | 2765.0 | 1.22 | 0.748 |

\* NOTE \* No outlier at the 5% level of significance



**Assumption 4, Independence** – Run order plot with runs of positive and negative residuals shown in figure below:



P-values for clustering and trends are above 0.05, therefore the data is random and stable with a confidence level of 95%.

***3.11.c*** *Use the Fisher LSD method with 𝛼 = 0.05 to make comparisons between pairs of means.*

The following output was produced using the data in Table 3.11.1 in Minitab:

**Figure 3.11.c.1 Minitab Residual Analysis**

**Fisher Pairwise Comparisons**

**Grouping Information Using the Fisher LSD Method and 95% Confidence**

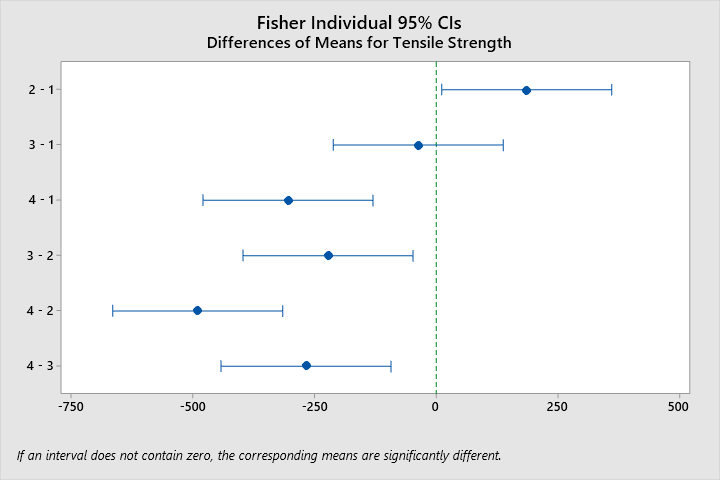
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Mixing Technique** | **N** | **Mean** | **Grouping** | | |
| 2 | 4 | 3156.3 | A |  |  |
| 1 | 4 | 2971.0 |  | B |  |
| 3 | 4 | 2933.8 |  | B |  |
| 4 | 4 | 2666.3 |  |  | C |

*Means that do not share a letter are significantly different.*

**Fisher Individual Tests for Differences of Means**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Difference of Levels** | **Difference of Means** | **SE of Difference** | **95% CI** | **T-Value** | **Adjusted P-Value** |
| 2 - 1 | 185.3 | 80.1 | (10.8, 359.7) | 2.31 | 0.039 |
| 3 - 1 | -37.3 | 80.1 | (-211.7, 137.2) | -0.47 | 0.650 |
| 4 - 1 | -304.8 | 80.1 | (-479.2, -130.3) | -3.81 | 0.003 |
| 3 - 2 | -222.5 | 80.1 | (-397.0, -48.0) | -2.78 | 0.017 |
| 4 - 2 | -490.0 | 80.1 | (-664.5, -315.5) | -6.12 | 0.000 |
| 4 - 3 | -267.5 | 80.1 | (-442.0, -93.0) | -3.34 | 0.006 |

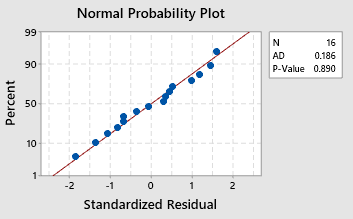
*Simultaneous confidence level = 81.57%*



The differences between (2-1), (4-1), (3-2), (4-2) and (4-2) have significant Fisher LSD Pairwise Comparison Adjusted P-values of 0.039, 0.003, 0.017, 0.000, and 0.006 respectively. Therefore, these combinations of factor means are not equivalent, with a confidence level of 95%.

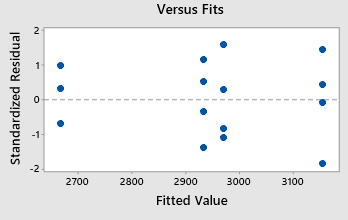
***3.11.d*** *Construct a normal probability plot of the residuals. What conclusion would you draw about the validity of the normality assumption?*

See Figure 3.11.a.2 Normal Probability plot. As stated in part 3.11.a,The p-value for the A.D. test for normality for this data is 0.890 (Not Significant), as seen near the Normal probability plot in the Minitab output below. Given a significance level of 0.05, this p-value will result in a failure to reject the null hypothesis that the data is normally distributed. In other words, the data is normally distributed, with a confidence level of 95%.



***3.11.e*** *Plot the residuals versus the predicted tensile strength. Comment on the plot.*

See Figure 3.11.a.2 versus fits plot. The lowest mean, mixing technique 4, at 2666.3, seems to be less widely dispersed as the other three, and significantly farther from any other means than the other means are. Mixing techniques 3 and 1 are in the middle and are much closer together. Mixing technique 2 has the highest mean tensile strength and average spread of values. None of the values shown are more than 2 standard deviations from the mean, supporting the assumption of no outliers.



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

**3. 33 Conditions**

|  |
| --- |
| *3.331 Add Levene’s Test, Tukey’s Multiple Comparison in SG3 Figure 8 order. Justify need for transformation in part (c).* |

**Table 3.33.1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Method | Count | | | | |
| 1 | 31 | 10 | 21 | 4 | 1 |
| 2 | 62 | 40 | 24 | 30 | 35 |
| 3 | 53 | 27 | 120 | 97 | 68 |

***3.33.a*** *Do all methods have the same effect on mean particle count?*

The data above was processed using the General Linear Model function using Minitab with the following results:

**Figure 3.33.a.1 Minitab ANOVA Table**

**General Linear Model: Count versus Method**

**Method**

|  |  |
| --- | --- |
| Factor coding | (-1, 0, +1) |

**Factor Information**

|  |  |  |  |
| --- | --- | --- | --- |
| **Factor** | **Type** | **Levels** | **Values** |
| Method | Fixed | 3 | 1, 2, 3 |

**Analysis of Variance**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Adj SS** | **Adj MS** | **F-Value** | **P-Value** |
| Method | 2 | 8964 | 4481.9 | 7.91 | 0.006 |
| Error | 12 | 6796 | 566.3 |  |  |
| Total | 14 | 15760 |  |  |  |

**Model Summary**

|  |  |  |  |
| --- | --- | --- | --- |
| **S** | **R-sq** | **R-sq(adj)** | **R-sq(pred)** |
| 23.7978 | 56.88% | 49.69% | 32.62% |

**Coefficients**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Term** | **Coef** | **SE Coef** | **T-Value** | **P-Value** | **VIF** |
| Constant | 41.53 | 6.14 | 6.76 | 0.000 |  |
| Method |  |  |  |  |  |
| 1 | -28.13 | 8.69 | -3.24 | 0.007 | 1.33 |
| 2 | -3.33 | 8.69 | -0.38 | 0.708 | 1.33 |

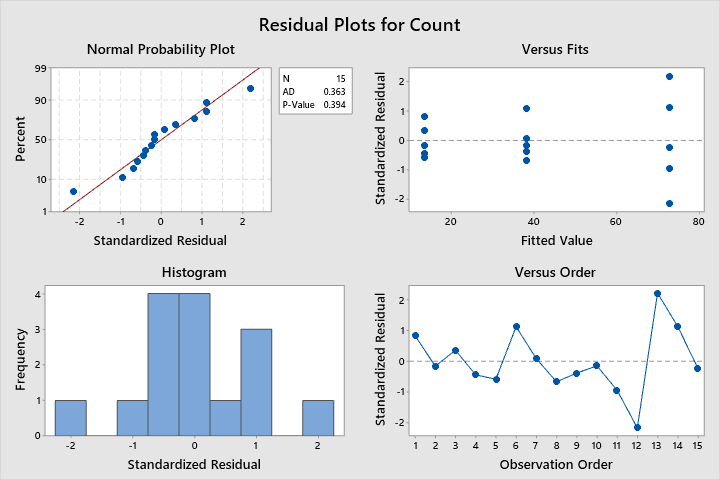
**Regression Equation**

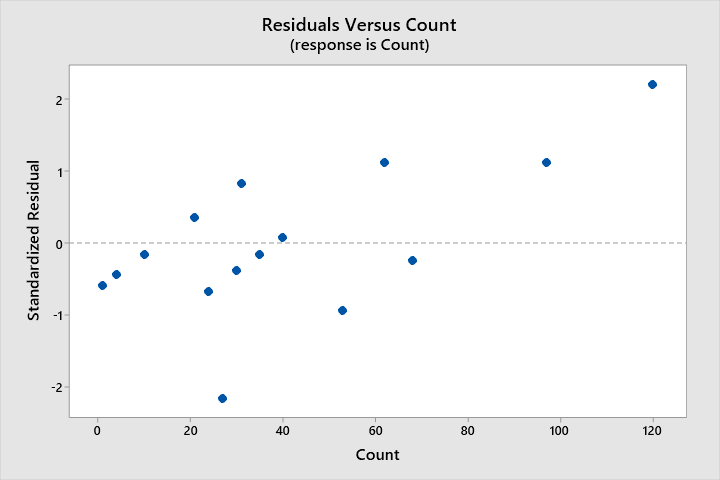
|  |  |  |
| --- | --- | --- |
| Count | = | 41.53 - 28.13 Method\_1 - 3.33 Method\_2 + 31.47 Method\_3 |

**Fits and Diagnostics for Unusual Observations**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Obs** | **Count** | **Fit** | **Resid** | **Std Resid** |  |
| 12 | 27.0 | 73.0 | -46.0 | -2.16 | R |
| 13 | 120.0 | 73.0 | 47.0 | 2.21 | R |

*R  Large residual*



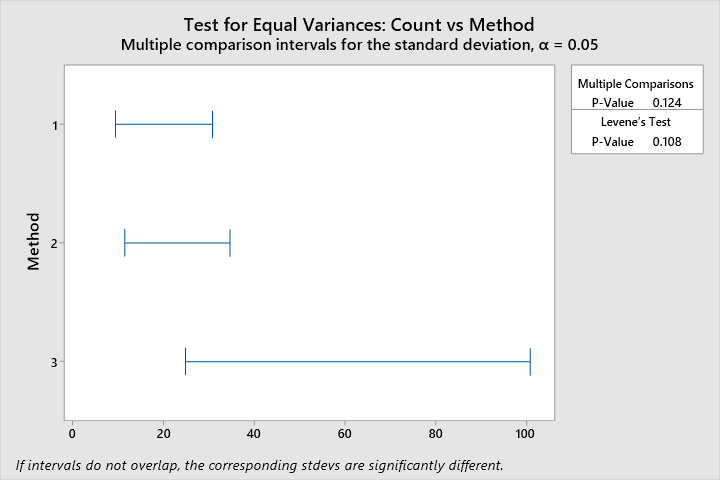


Given a P-value of 0.006 from the ANOVA table above, and given a significance level of 0.05, the null hypothesis ought to be rejected. Therefore, it is unlikely that the means for the 4 groups in table 3.11.1 are equivalent. It can be said with 95% certainty that at least one of the means is not equivalent to the others. This conclusion is subject to the validity of the assumptions, which need to be verified.

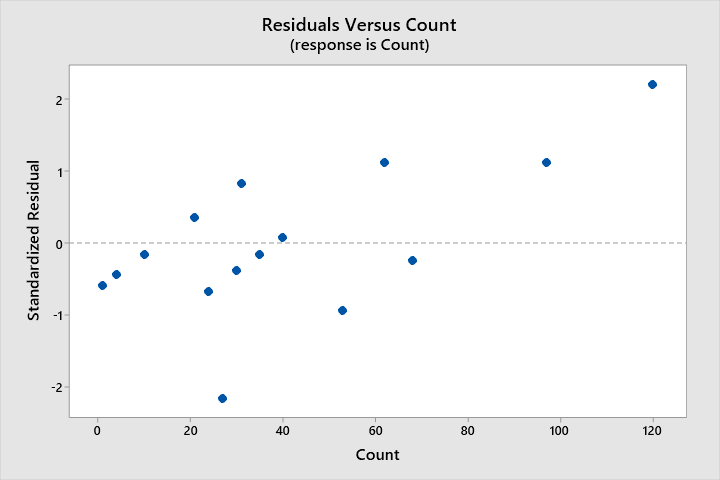
It is important to check the assumptions for any statistical analysis. Page 16 of Study Guide 3 for this course contains a “Summary of Model Adequacy Checking” which will be utilized as an outline to ensure all assumptions are addressed properly (This satisfies condition 3.113).

**Assumption 1, Normality** – The p-value for the A.D. test for normality for this data is 0.363 (Not Significant), as seen near the Normal probability plot in the Minitab output in Figure 3.33.a.1. Given a significance level of 0.05, this p-value will result in a failure to reject the null hypothesis that the data is normally distributed. In other words, the data is normally distributed, with a confidence level of 95%.

**Assumption 2, Constant Variance** – Levene’s Test can be utilized to verify this assumption.



With an insignificant p-value of 0.108, and a significance level of 0.05, it is likely that the variances are equal for the 3 groups. This conclusion is questionable, given the low p-value and the large discrepancy of the size of the variation between methods 1/2 and method 3. The residual/predicted plot is below.



There appears to be a “funnel” shape to the residual versus count data. This is indicative of increasing variation, which is undesirable.

**Assumption 3, No Outliers** – The data was checked for outliers using Minitab. There are no outliers in this data. See Minitab output below:

**Outlier Test: Count**

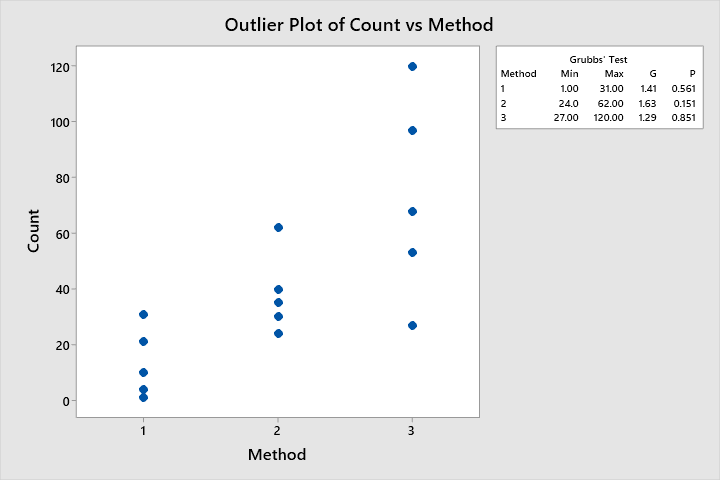
**Method**

|  |  |
| --- | --- |
| Null hypothesis | All data values come from the same normal population |
| Alternative hypothesis | Smallest or largest data value is an outlier |
| Significance level | α = 0.05 |

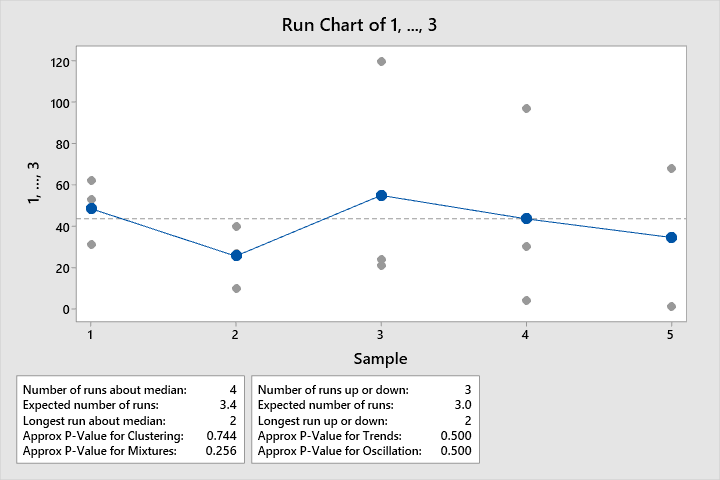
**Grubbs' Test**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Method** | **N** | **Mean** | **StDev** | **Min** | **Max** | **G** | **P** |
| Count | 1 | 5 | 13.40 | 12.46 | 1.00 | 31.00 | 1.41 | 0.561 |
|  | 2 | 5 | 38.20 | 14.57 | 24.00 | 62.00 | 1.63 | 0.151 |
|  | 3 | 5 | 73.0 | 36.5 | 27.0 | 120.0 | 1.29 | 0.851 |

\* NOTE \* No outlier at the 5% level of significance



**Assumption 4, Independence** – Run order plot with runs of positive and negative residuals shown in figure below:



P-values for clustering and trends are above 0.05, therefore the data is random and stable with a confidence level of 95%. This assumption relies on that the data were recorded in run order, otherwise the independence test cannot be performed.

***3.33.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?*

Referring to the conclusions drawn in 3.33.a **Assumption 2,** there appears to be a funnel shape in the Standardized residual versus predicted value plot, which is indicative of increasing variation, which is undesirable. An effective countermeasure to this situation is to transform the data using a sqrt(y) or ln(y) transformation.

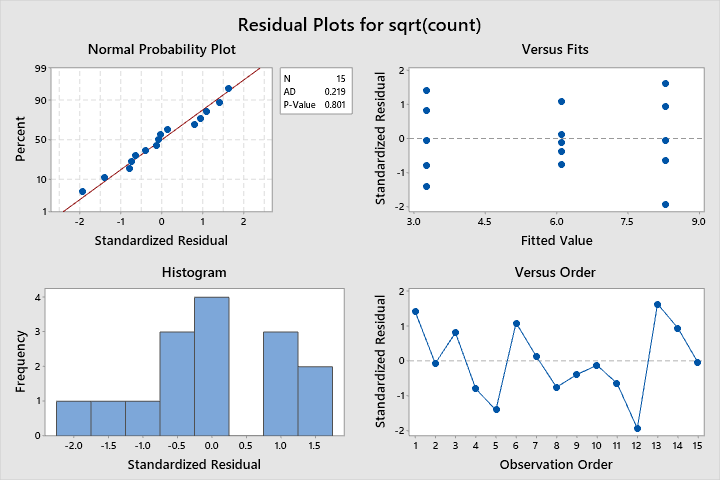
***3.33.c*** *Based on your answer to part (b), conduct another analysis of the particle count data and draw appropriate conclusions.*

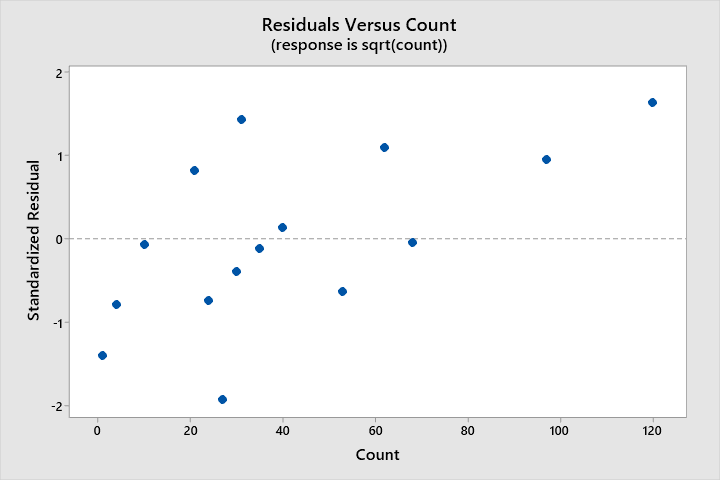
The data was transformed to sqrt(count) and reprocessed using Minitab. Some results are below:

**Figure 3.33.c.1 Minitab ANOVA Table of sqrt(Count) versus Method**

**Analysis of Variance**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Adj SS** | **Adj MS** | **F-Value** | **P-Value** |
| Method | 2 | 63.90 | 31.950 | 9.84 | 0.003 |
| Error | 12 | 38.96 | 3.247 |  |  |
| Total | 14 | 102.86 |  |  |  |





With a p-value of 0.003, the conclusion of the ANOVA is to reject the null hypothesis that all means for the 3 methods are equal. The means are different, with a confidence level of 95%.

The F-test is significant; therefore it is appropriate to perform a pairwise comparison of means using Tukey’s procedure:

**Figure 3.33.c.1 Tukey Pairwise Comparison Output using Minitab**

**Comparisons for sqrt(count)**

**Tukey Pairwise Comparisons: Method**

**Grouping Information Using the Tukey Method and 95% Confidence**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Method** | **N** | **Mean** | **Grouping** | |
| 3 | 5 | 8.30516 | A |  |
| 2 | 5 | 6.09817 | A | B |
| 1 | 5 | 3.26252 |  | B |

*Means that do not share a letter are significantly different.*

**Tukey Simultaneous Tests for Differences of Means**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Difference of Method Levels** | **Difference of Means** | **SE of Difference** | **Simultaneous 95% CI** | **T-Value** | **Adjusted P-Value** |
| 2 - 1 | 2.84 | 1.14 | (-0.20, 5.87) | 2.49 | 0.068 |
| 3 - 1 | 5.04 | 1.14 | (2.00, 8.08) | 4.42 | 0.002 |
| 3 - 2 | 2.21 | 1.14 | (-0.83, 5.25) | 1.94 | 0.171 |

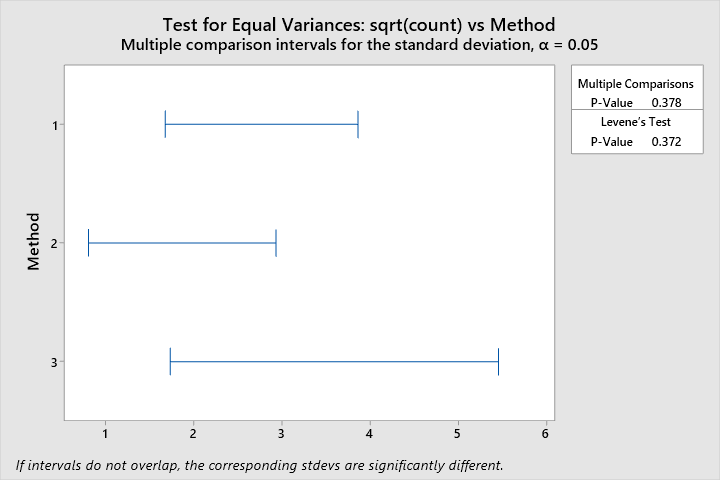
*Individual confidence level = 97.94%*

Given a significance level of 0.05, the means for methods 3 and 1 are different, according to the Tukey Pairwise Comparisons method output above.

The assumptions need to be verified for the transformed data ANOVA.

**Assumption 1, Normality** – The p-value for the A.D. test for normality for this data is 0.801 (Not Significant), as seen near the Normal probability plot in the Minitab output above. Given a significance level of 0.05, this p-value will result in a failure to reject the null hypothesis that the data is normally distributed. In other words, the data is normally distributed, with a confidence level of 95%.

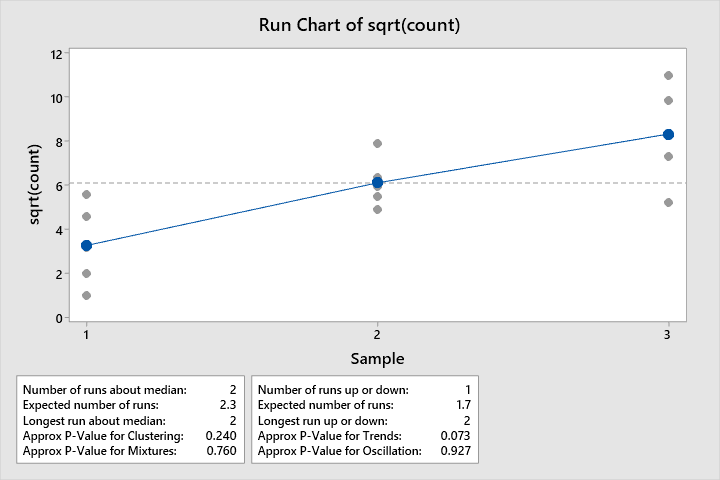
**Assumption 2, Constant Variance** – Levene’s Test can be utilized to verify this assumption. With a P-Value of 0.372, the variances of the groups of methods are equal.



After the square root transformation, the “funnel” shape to the residual versus count data is reduced. The transformation has reduced the appearance of increasing variation.

**Assumption 3, No Outliers** – There are no values more than 2 standard deviations from the mean, as shown in **Figure 3.33.c.1 versus fits plot.**

**Assumption 4, Independence** – Run order plot with runs of positive and negative residuals shown in figure below:



P-values for clustering and trends are above 0.05, therefore the data is random and stable with a confidence level of 95%. This assumption relies on that the data were recorded in run order, otherwise the independence test cannot be performed.

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

***3.61 Conditions*** *(The bar below cannot be deleted.. strange)*

|  |
| --- |
| 3.614 Example 3.1 and 3.6. Replace largest 725 by 250 and compare KW and standard methods. |

**Table 3.61.1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| RF power (W) | Observed Etch Rate Å/min | | | | |
| 160 | 575 | 542 | 530 | 539 | 570 |
| 180 | 565 | 593 | 590 | 579 | 610 |
| 200 | 600 | 651 | 610 | 637 | 629 |
| 220 | ~~725~~ 250 | 700 | 715 | 685 | 710 |

The data were processed using the general linear model function in Minitab. The results are below:

**General Linear Model: Etch Rate versus RF Power**

**Method**

|  |  |
| --- | --- |
| Factor coding | (-1, 0, +1) |

**Factor Information**

|  |  |  |  |
| --- | --- | --- | --- |
| **Factor** | **Type** | **Levels** | **Values** |
| RF Power | Fixed | 4 | 160, 180, 200, 220 |

**Analysis of Variance**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Adj SS** | **Adj MS** | **F-Value** | **P-Value** |
| RF Power | 3 | 15927 | 5309 | 0.50 | 0.685 |
| Error | 16 | 168739 | 10546 |  |  |
| Total | 19 | 184666 |  |  |  |

**Model Summary**

|  |  |  |  |
| --- | --- | --- | --- |
| **S** | **R-sq** | **R-sq(adj)** | **R-sq(pred)** |
| 102.695 | 8.62% | 0.00% | 0.00% |

**Coefficients**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Term** | **Coef** | **SE Coef** | **T-Value** | **P-Value** | **VIF** |
| Constant | 594.0 | 23.0 | 25.87 | 0.000 |  |
| RF Power |  |  |  |  |  |
| 160 | -42.8 | 39.8 | -1.08 | 0.298 | 1.50 |
| 180 | -6.6 | 39.8 | -0.17 | 0.870 | 1.50 |
| 200 | 31.4 | 39.8 | 0.79 | 0.441 | 1.50 |

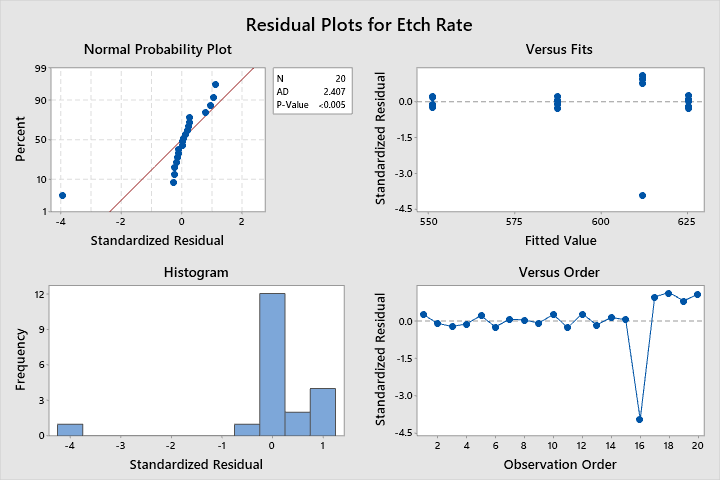
**Regression Equation**

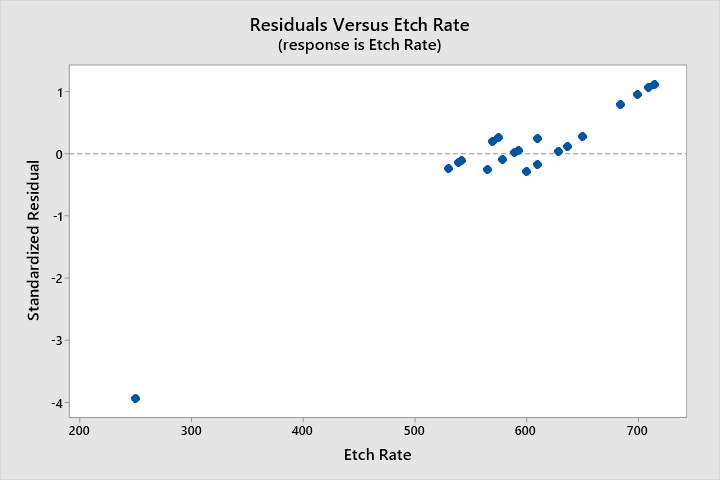
|  |  |  |
| --- | --- | --- |
| Etch Rate | = | 594.0 - 42.8 RF Power\_160 - 6.6 RF Power\_180 + 31.4 RF Power\_200 + 18.0 RF Power\_220 |

**Fits and Diagnostics for Unusual Observations**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Obs** | **Etch Rate** | **Fit** | **Resid** | **Std Resid** |  |
| 16 | 250.0 | 612.0 | -362.0 | -3.94 | R |

*R  Large residual*





*What effect does this have on the usual analysis of variance?*

The introduction of an outlier due to a mistaken record has resulted in a violation of multiple assumptions necessary to accept the conclusions of the ANOVA test. In particular, assumption 1 (Normality) cannot be verified, as the p-value of the A.D. test is less than 0.05. Assumption 3 (Outlier Analysis) is also violated, given the value 250 is beyond 3 standard deviations from the expected value. Upon recognition of the violation of the assumptions due to an extreme outlier, Kruskal-Wallis nonparametric ANOVA can be performed.

*What effect does it have on the Kruskal–Wallis test?*

The following are Minitab Kruskal-Wallis outputs for the correct vs. altered data:

**Kruskal-Wallis Test: Etch Rate versus RF Power (Altered Data)**

**Descriptive Statistics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **RF Power** | **N** | **Median** | **Mean Rank** | **Z-Value** |
| 160 | 5 | 542 | 4.4 | -2.66 |
| 180 | 5 | 590 | 8.9 | -0.70 |
| 200 | 5 | 629 | 13.7 | 1.40 |
| 220 | 5 | 700 | 15.0 | 1.96 |
| Overall | 20 |  | 10.5 |  |

**Test**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Null hypothesis | | H₀: All medians are equal | | | |
| Alternative hypothesis | | H₁: At least one median is different | | | |
| **Method** | **DF** | | **H-Value** | **P-Value** |
| Not adjusted for ties | 3 | | 10.04 | 0.018 |
| Adjusted for ties | 3 | | 10.04 | 0.018 |

**Kruskal-Wallis Test: Etch Rate versus RF Power (Correct Data)**

**Descriptive Statistics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **RF Power** | **N** | **Median** | **Mean Rank** | **Z-Value** |
| 160 | 5 | 542 | 3.4 | -3.10 |
| 180 | 5 | 590 | 7.9 | -1.13 |
| 200 | 5 | 629 | 12.7 | 0.96 |
| 220 | 5 | 710 | 18.0 | 3.27 |
| Overall | 20 |  | 10.5 |  |

**Test**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Null hypothesis | | H₀: All medians are equal | | | |
| Alternative hypothesis | | H₁: At least one median is different | | | |
| **Method** | **DF** | | **H-Value** | **P-Value** |
| Not adjusted for ties | 3 | | 16.89 | 0.001 |
| Adjusted for ties | 3 | | 16.91 | 0.001 |

The p-value is significant in both cases. The introduction of an erroneous observation does not change the underlying conclusion of the K-W test, but the K-W test does require verification of equal variances, which must be tested before the conclusion of the test can be fully accepted.