STAT 8210 – Applied Experimental Design

Module 2 Homework

Due 1/17/2020

Connor Armstrong

*2.28 The following are the burning times (in minutes) of chemical flares of two different formulations. The design engineers are interested in both the mean and variance of the burning times.*

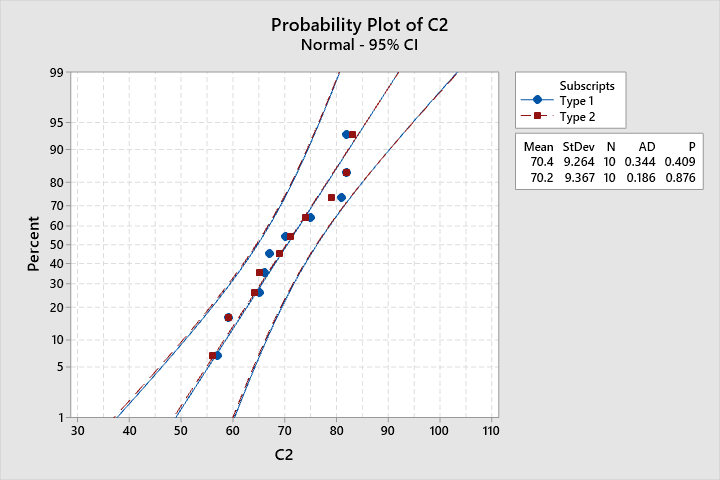
Table 2.28.1 Burning Times (in minutes) of Chemical Flares of 2 Different Formulations

|  |  |  |  |
| --- | --- | --- | --- |
| **Type 1** | | **Type 2** | |
| 65 | 82 | 64 | 56 |
| 81 | 67 | 71 | 69 |
| 57 | 59 | 83 | 74 |
| 66 | 75 | 59 | 82 |
| 82 | 70 | 65 | 79 |

*(a) Test the hypothesis that the two variances are equal. Use α = 0.05.*

The data were input into Minitab and the two columns were stacked in a new worksheet. The next step is to test for normality. The results are below.

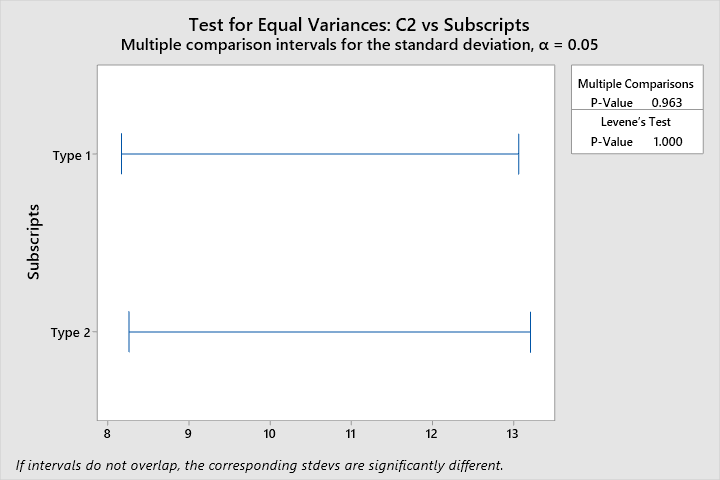
Figure 2.28.a.1 Evaluating Normality Assumption for Question 2.28.a



Given the P-values of 0.409 and 0.876 are greater than the significance level of 0.05, it is likely that the populations from which these samples were drawn are normal.

Minitab is used to implement an ANOVA test for equal variance for the data provided, shown below.

Figure 2.28.a.2 Evaluating Homogeneity of Variance for Question 2.28.a



The hypothesis of equal variance is not significant (P=1.000). Thus, it is very likely that the variances are equal for the two groups.

*(b) Using the results of (a), test the hypothesis that the mean burning times are equal. Use α = 0.05. What is the P-value for this test?*

The previous part of this question addressed the 2 major assumptions necessary to accept the Two-Sample t-test. A Two-Sample t-test was computed in Minitab with results shown below.

Figure 2.28.b.1 Minitab Output for Two-Sample T-Test Addressing Question 2.28.b

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 28 SEP COL  **Two-Sample T-Test and CI: Type 1, Type 2**  **Method**   |  | | --- | | μ₁: mean of Type 1 | | µ₂: mean of Type 2 | | Difference: μ₁ - µ₂ |   *Equal variances are assumed for this analysis.*  **Descriptive Statistics**   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Sample** | **N** | **Mean** | **StDev** | **SE Mean** | | Type 1 | 10 | 70.40 | 9.26 | 2.9 | | Type 2 | 10 | 70.20 | 9.37 | 3.0 |   **Estimation for Difference**   |  |  |  | | --- | --- | --- | | **Difference** | **Pooled StDev** | **95% CI for Difference** | | 0.20 | 9.32 | (-8.55, 8.95) |   **Test**   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Null hypothesis | | | H₀: μ₁ - µ₂ = 0 | | | Alternative hypothesis | | | H₁: μ₁ - µ₂ ≠ 0 | | | **T-Value** | **DF** | **P-Value** | | | 0.05 | 18 | 0.962 | | |

The P-Value is not significant (P=0.962). Therefore, there is not sufficient evidence to reject the null hypothesis in this case. It is likely that the means are equal, with a confidence level of 95%.

*(c) Discuss the role of the normality assumption in this problem. Check the assumption of normality for both types of flares.*

If the normality assumption was violated, then it would be appropriate to implement a different type of t-test to answer this question (Montgomery, DOE 8th Ed. Page 42). When assumptions are violated, the performance and accuracy of the t-test will be affected. To satisfy the normality assumption, an Anderson-Darling Normality Test was performed on the dataset.

Referring to Figure 2.28.a.1, the P-values of 0.409 and 0.876 for Type 1 and Type 2 respectively, are greater than the significance level of 0.05. Therefore, it is likely that the population from which these samples were drawn is normal.

*2.34 The diameter of a ball bearing was measured by 12 inspectors, each using two different kinds of calipers. The results are as follows:*

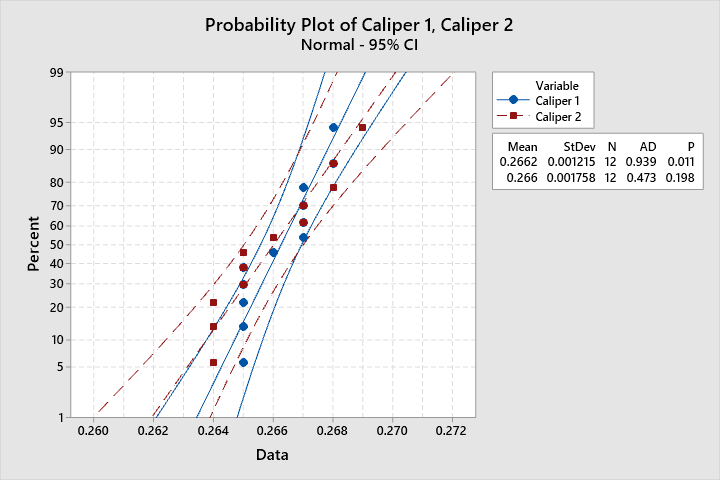
Table 2.34.1 Ball Bearing Diameter Using 2 Kinds of Calipers

|  |  |  |
| --- | --- | --- |
| **Inspector** | **Caliper 1** | **Caliper 2** |
| 1 | 0.265 | 0.264 |
| 2 | 0.265 | 0.265 |
| 3 | 0.266 | 0.264 |
| 4 | 0.267 | 0.266 |
| 5 | 0.267 | 0.267 |
| 6 | 0.265 | 0.268 |
| 7 | 0.267 | 0.264 |
| 8 | 0.267 | 0.265 |
| 9 | 0.265 | 0.265 |
| 10 | 0.268 | 0.267 |
| 11 | 0.268 | 0.268 |
| 12 | 0.265 | 0.269 |

*(a) Is there a significant difference between the means of the population of measurements from which the two samples were selected? Use α = 0.05.*

A two-sample t-test will be computed in order to address this question. It is desirable to check the assumptions for this analysis. The 2 assumptions to be checked are that the data are normally distributed and that variances for the two groups are equal. The following Minitab and SAS outputs address these assumptions.

Figure 2.34.a.1 Minitab Probability Plot for Validation of Normality Assumption



Given a significant P-value of 0.011 (less than the significance level of 0.05), it is likely that the population from which the samples in group “Caliper 1” were drawn is not normal.

Table 2.34.a.1 SAS Testing of t-Test Assumptions for Question 2.34.a

| **Goodness-of-Fit Tests for Normal Distribution** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Kolmogorov-Smirnov** | **D** | 0.20898559 | **Pr > D** | <0.010 |
| **Cramer-von Mises** | **W-Sq** | 0.13637989 | **Pr > W-Sq** | 0.035 |
| **Anderson-Darling** | **A-Sq** | 0.77966901 | **Pr > A-Sq** | 0.039 |

The highlighted P-value for the Anderson Darling test (P=0.039 is less than α=0.05) suggests that the data is not normal. The equal variance assumption need not be checked.

Having violated the assumption for normality, it is appropriate to use the Mann-Whitney test, otherwise known as the Wilcoxon Rank-Sum Test.

Figure 2.34.a.1 Minitab Mann-Whitney Test for Question 2.34.a

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 34 SEP COL  **Mann-Whitney: Caliper 1, Caliper 2**  **Method**   |  | | --- | | η₁: median of Caliper 1 | | η₂: median of Caliper 2 | | Difference: η₁ - η₂ |   **Descriptive Statistics**   |  |  |  | | --- | --- | --- | | **Sample** | **N** | **Median** | | Caliper 1 | 12 | 0.2665 | | Caliper 2 | 12 | 0.2655 |   **Estimation for Difference**   |  |  |  | | --- | --- | --- | | **Difference** | **CI for Difference** | **Achieved Confidence** | | 0.0000000 | (-0.0010000, 0.0020000) | 95.36% |   **Test**   |  |  |  |  | | --- | --- | --- | --- | | Null hypothesis | | H₀: η₁ - η₂ = 0 | | | Alternative hypothesis | | H₁: η₁ - η₂ ≠ 0 | | | **Method** | **W-Value** | | **P-Value** | | | Not adjusted for ties | 159.00 | | 0.624 | | | Adjusted for ties | 159.00 | | 0.613 | | |

Similarly, the Mann-Whitney test will be computed in SAS:

ods rtf;

ods graphics on;

**proc** **npar1way** data =study2.hw234 wilcoxon;

class group;

var meas;

**run**;

ods graphics off;

ods rtf close;

See output in the following page.

Figure 2.34.a.2 SAS Mann-Whitney Test for Question 2.34.a

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | **Wilcoxon Two-Sample Test** | | | --- | --- | | **Statistic** | 159.0000 | |  |  | | **Normal Approximation** |  | | **Z** | 0.5057 | | **One-Sided Pr > Z** | 0.3065 | | **Two-Sided Pr > |Z|** | 0.6131 | |  |  | | **t Approximation** |  | | **One-Sided Pr > Z** | 0.3089 | | **Two-Sided Pr > |Z|** | 0.6179 | | **Z includes a continuity correction of 0.5.** | |  | **Kruskal-Wallis Test** | | | --- | --- | | **Chi-Square** | 0.2867 | | **DF** | 1 | | **Pr > Chi-Square** | 0.5923 | |

The P-Value of the analysis performed is not significant (PMINITAB=0.624 and PSAS=0.613 are greater than α=0.05). Therefore, there is not sufficient evidence to reject the null hypothesis that η₁ - η₂ = 0. Subsequently, it is likely that the true value of the medians for the two groups of measurements are equal. The discrepancy between the P-Values for the two softwares is unknown.

*(b) Find the P-value for the test in part (a).*

The P-Values for the Mann-Whitney Test in both SAS and Minitab are appropriately designated.

*(c) Construct a 95 percent confidence interval on the difference in mean diameter measurements for the two types of calipers.*

A nonparametric approach is necessary, given the normality assumption is in question. Referring to the study guide, it is found on page 17 that Table 12 contains the “nonparametric equivalent, the Wilcoxon Signed-Rank Test” of Table 11, which contains a 95% CI for the mean difference with a paired T-Test. There is some contradiction here, between the Table numbers and their titles. This has resulted in substantial confusion in this question.

The following code was ran in SAS to construct the 95% CI for the difference in mean diameter for the two types of calipers:

ods rtf;

ods graphics on;

**proc** **ttest** data =study2.hw234b;

paired caliper1\*caliper2;

**run**;

ods graphics off;

ods rtf close;

Figure 2.34.c.1 SAS Paired T-Test for Question 2.34.c

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | **Mean** | **95% CL Mean** | | **Std Dev** | **95% CL Std Dev** | | | --- | --- | --- | --- | --- | --- | | 0.000250 | -0.00102 | 0.00152 | 0.00201 | 0.00142 | 0.00341 |  | **DF** | **t Value** | **Pr > |t|** | | --- | --- | --- | | 11 | 0.43 | 0.6742 | |

While table 11.b in the study guide is titled “Minitab Wilcoxon Signed-Rank Test for Example 2,” it does not provide sufficient information to construct the requested confidence interval. Also, there does not seem to be an appropriate nonparametric method for determining the confidence interval in Minitab within the study guide or text for this chapter.

The 95% CI on the difference in mean diameter measurements between the two types of calipers is (-0.00102, 0.00152).

*2.37 The deflection temperature under load for two different formulations of ABS plastic pipe is being studied. Two samples of 12 observations each are prepared using each formulation and the deflection temperatures (in degrees F) are reported below:*

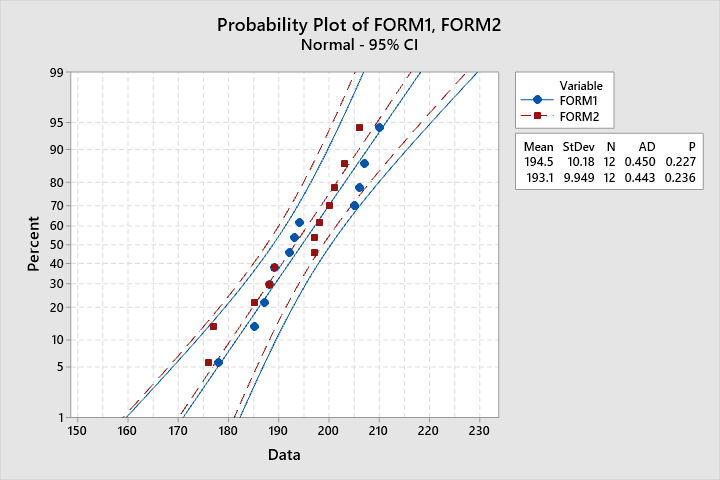
Table 2.37.1 Deflection Temperature Under Load for Two Formulations

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Formulation 1** | | | **Formulation 2** | | |
| 206 | 193 | 192 | 177 | 176 | 198 |
| 188 | 207 | 210 | 197 | 185 | 188 |
| 205 | 185 | 194 | 206 | 200 | 189 |
| 187 | 189 | 178 | 201 | 197 | 203 |

*(a) Construct normal probability plots for both samples. Do these plots support assumptions of normality and equal variance for both samples?*

The following probability plots were generated using Minitab and SAS:

Figure 2.37.a.1 Minitab Probability Plot for Question 2.37.a



Upon close inspection of Figure 2.37.a.1, the confidence and prediction intervals for the probability plots for the two formulations have significant overlap, and very similar slope. These factors are good indicators that the assumptions for equal variance and normality are likely satisfied. This will be explored further.

Figure 2.37.a.2 Minitab Test for Homogeneity of Variance for Question 2.37.a

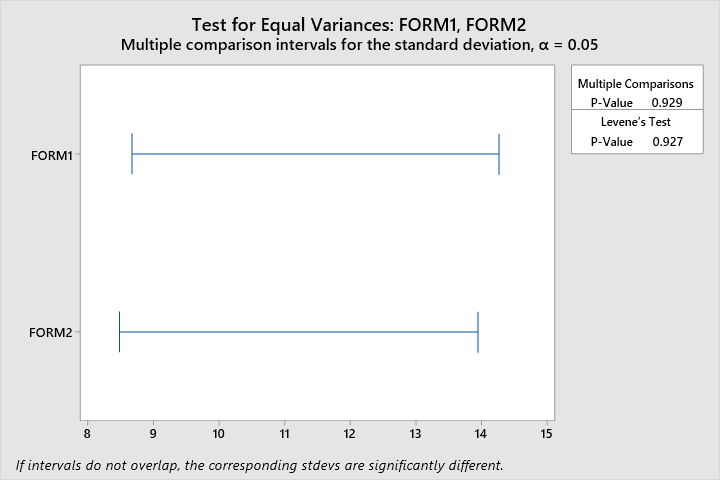


Figure 2.37.a.2 supports the conclusion that the variances for the 2 formulations are equal.

Figure 2.37.a.2 SAS Probability Plot for Question 2.37.a



*The above plot is meant to partially satisfy the requirement to “+Use both SAS and Minitab.”*

Table 2.37.a.1 SAS Output Goodness of Fit for Normal Distribution for Question 2.37.a

| **Goodness-of-Fit Tests for Normal Distribution** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Kolmogorov-Smirnov** | **D** | 0.11300067 | **Pr > D** | >0.150 |
| **Cramer-von Mises** | **W-Sq** | 0.05124138 | **Pr > W-Sq** | >0.250 |
| **Anderson-Darling** | **A-Sq** | 0.36188581 | **Pr > A-Sq** | >0.250 |

Table 2.37.a.2 SAS Output Equal Variance Test for Question 2.37.a

| **Levene's Test for Homogeneity of data Variance ANOVA of Squared Deviations from Group Means** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **form** | 1 | 104.5 | 104.5 | 0.01 | 0.9145 |
| **Error** | 22 | 195107 | 8868.5 |  |  |

The above tables further verify the conclusions reached previously, that the two formulations likely come from normal populations with equal variance, with a confidence level of 95%.

*(b) Do the data support the claim that the mean deflection temperature under load for formulation 1 exceeds that of formulation 2? Use α = 0.05.*

In order to determine whether the mean temperature for formulation 1 exceeds that of formulation 2, a two-sample t-test will be performed. Unlike the previous examples, this question is interested in one temperature being greater than another, rather than just not equal. Thusly, a one sided test will be performed.

Figure 2.37.b.1 Minitab Two Sample T-Test for Question 2.37.a

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 37 SEP COL  **Two-Sample T-Test and CI: FORM1, FORM2**  **Method**   |  | | --- | | μ₁: mean of FORM1 | | µ₂: mean of FORM2 | | Difference: μ₁ - µ₂ |   *Equal variances are assumed for this analysis.*  **Descriptive Statistics**   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Sample** | **N** | **Mean** | **StDev** | **SE Mean** | | FORM1 | 12 | 194.5 | 10.2 | 2.9 | | FORM2 | 12 | 193.08 | 9.95 | 2.9 |   **Estimation for Difference**   |  |  |  | | --- | --- | --- | | **Difference** | **Pooled StDev** | **95% Lower Bound for Difference** | | 1.42 | 10.06 | -5.64 |   **Test**   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Null hypothesis | | | H₀: μ₁ - µ₂ = 0 | | | Alternative hypothesis | | | H₁: μ₁ - µ₂ > 0 | | | **T-Value** | **DF** | **P-Value** | | | 0.34 | 22 | 0.367 | | |

Given a non-significant P-Value of 0.367 to the Two-Sample T-Test, there is not enough evidence to reject the null hypothesis that the means of the two formulations are equivalent. In other words, with 95% certainty, the mean deflection temperature of formulation 1 does not exceed that of formulation 2.

*(c) What is the P-value for the test in part (a)?*

The P-Values for the two Anderson-Darling tests completed with respect to the two formulations are P1 = 0.227 and P2 = 0.236, as seen in Figure 2.37.a.1.