CSC 423 Project 1

Part A:

1:

| **Obs** | **measurer** | **blue** | **white** |
| --- | --- | --- | --- |
| **1** | 1 | 0.108 | 0.109 |
| **2** | 2 | 0.101 | 0.097 |
| **3** | 3 | 0.104 | 0.100 |
| **4** | 4 | 0.105 | 0.095 |
| **5** | 5 | 0.105 | 0.103 |
| **6** | 6 | 0.108 | 0.103 |
| **7** | 7 | 0.106 | 0.095 |
| **8** | 8 | 0.104 | 0.097 |
| **9** | 9 | 0.105 | 0.101 |
| **10** | 10 | 0.108 | 0.103 |
| **11** | 11 | 0.103 | 0.106 |
| **12** | 12 | 0.103 | 0.102 |
| **13** | 13 | 0.110 | 0.104 |
| **14** | 14 | 0.109 | 0.102 |
| **15** | 15 | 0.105 | 0.099 |
| **16** | 16 | 0.107 | 0.102 |
| **17** | 17 | 0.102 | 0.094 |
| **18** | 18 | 0.104 | 0.097 |
| **19** | 19 | 0.104 | 0.104 |
| **20** | 20 | 0.100 | 0.097 |

2:

| **Obs** | **measurer** | **color** | **thickness** |
| --- | --- | --- | --- |
| **1** | 1 | blue | 0.108 |
| **2** | 2 | blue | 0.101 |
| **3** | 3 | blue | 0.104 |
| **4** | 4 | blue | 0.105 |
| **5** | 5 | blue | 0.105 |
| **6** | 6 | blue | 0.108 |
| **7** | 7 | blue | 0.106 |
| **8** | 8 | blue | 0.104 |
| **9** | 9 | blue | 0.105 |
| **10** | 10 | blue | 0.108 |
| **11** | 11 | blue | 0.103 |
| **12** | 12 | blue | 0.103 |
| **13** | 13 | blue | 0.110 |
| **14** | 14 | blue | 0.109 |
| **15** | 15 | blue | 0.105 |
| **16** | 16 | blue | 0.107 |
| **17** | 17 | blue | 0.102 |
| **18** | 18 | blue | 0.104 |
| **19** | 19 | blue | 0.104 |
| **20** | 20 | blue | 0.100 |
| **21** | 1 | white | 0.109 |
| **22** | 2 | white | 0.097 |
| **23** | 3 | white | 0.100 |
| **24** | 4 | white | 0.095 |
| **25** | 5 | white | 0.103 |
| **26** | 6 | white | 0.103 |
| **27** | 7 | white | 0.095 |
| **28** | 8 | white | 0.097 |
| **29** | 9 | white | 0.101 |
| **30** | 10 | white | 0.103 |
| **31** | 11 | white | 0.106 |
| **32** | 12 | white | 0.102 |
| **33** | 13 | white | 0.104 |
| **34** | 14 | white | 0.102 |
| **35** | 15 | white | 0.099 |
| **36** | 16 | white | 0.102 |
| **37** | 17 | white | 0.094 |
| **38** | 18 | white | 0.097 |
| **39** | 19 | white | 0.104 |
| **40** | 20 | white | 0.097 |

3:

**The UNIVARIATE Procedure**

**Variable: blue**

| **Moments** | | | |
| --- | --- | --- | --- |
| **N** | 20 | **Sum Weights** | 20 |
| **Mean** | 0.10505 | **Sum Observations** | 2.101 |
| **Std Deviation** | 0.00266508 | **Variance** | 7.10263E-6 |
| **Skewness** | 0.07029943 | **Kurtosis** | -0.4856352 |
| **Uncorrected SS** | 0.220845 | **Corrected SS** | 0.00013495 |
| **Coeff Variation** | 2.53695981 | **Std Error Mean** | 0.00059593 |

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 0.105050 | **Std Deviation** | 0.00267 |
| **Median** | 0.105000 | **Variance** | 7.10263E-6 |
| **Mode** | 0.104000 | **Range** | 0.01000 |
|  |  | **Interquartile Range** | 0.00400 |

**Note: The mode displayed is the smallest of 2 modes with a count of 4.**

| **Tests for Location: Mu0=0** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Student's t** | **t** | 176.2793 | **Pr > |t|** | <.0001 |
| **Sign** | **M** | 10 | **Pr >= |M|** | <.0001 |
| **Signed Rank** | **S** | 105 | **Pr >= |S|** | <.0001 |

| **Quantiles (Definition 5)** | |
| --- | --- |
| **Level** | **Quantile** |
| **100% Max** | 0.1100 |
| **99%** | 0.1100 |
| **95%** | 0.1095 |
| **90%** | 0.1085 |
| **75% Q3** | 0.1075 |
| **50% Median** | 0.1050 |
| **25% Q1** | 0.1035 |
| **10%** | 0.1015 |
| **5%** | 0.1005 |
| **1%** | 0.1000 |
| **0% Min** | 0.1000 |

**The UNIVARIATE Procedure**

**Variable: white**

| **Moments** | | | |
| --- | --- | --- | --- |
| **N** | 20 | **Sum Weights** | 20 |
| **Mean** | 0.1005 | **Sum Observations** | 2.01 |
| **Std Deviation** | 0.00401969 | **Variance** | 0.00001616 |
| **Skewness** | 0.12695384 | **Kurtosis** | -0.5541375 |
| **Uncorrected SS** | 0.202312 | **Corrected SS** | 0.000307 |
| **Coeff Variation** | 3.99968994 | **Std Error Mean** | 0.00089883 |

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 0.100500 | **Std Deviation** | 0.00402 |
| **Median** | 0.101500 | **Variance** | 0.0000162 |
| **Mode** | 0.097000 | **Range** | 0.01500 |
|  |  | **Interquartile Range** | 0.00600 |

| **Tests for Location: Mu0=0** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Student's t** | **t** | 111.8121 | **Pr > |t|** | <.0001 |
| **Sign** | **M** | 10 | **Pr >= |M|** | <.0001 |
| **Signed Rank** | **S** | 105 | **Pr >= |S|** | <.0001 |

| **Quantiles (Definition 5)** | |
| --- | --- |
| **Level** | **Quantile** |
| **100% Max** | 0.1090 |
| **99%** | 0.1090 |
| **95%** | 0.1075 |
| **90%** | 0.1050 |
| **75% Q3** | 0.1030 |
| **50% Median** | 0.1015 |
| **25% Q1** | 0.0970 |
| **10%** | 0.0950 |
| **5%** | 0.0945 |
| **1%** | 0.0940 |
| **0% Min** | 0.0940 |

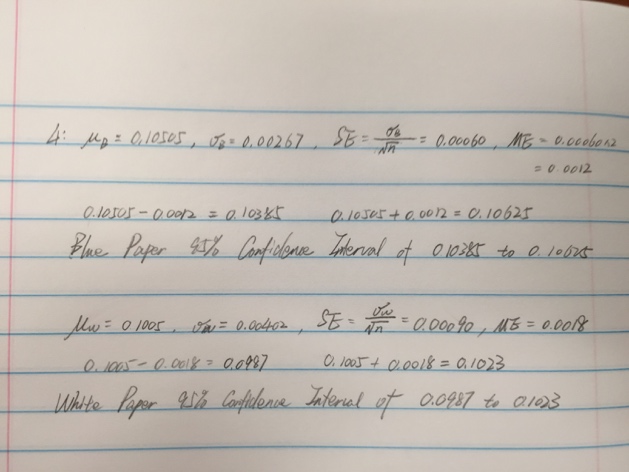
4:

By SAS:

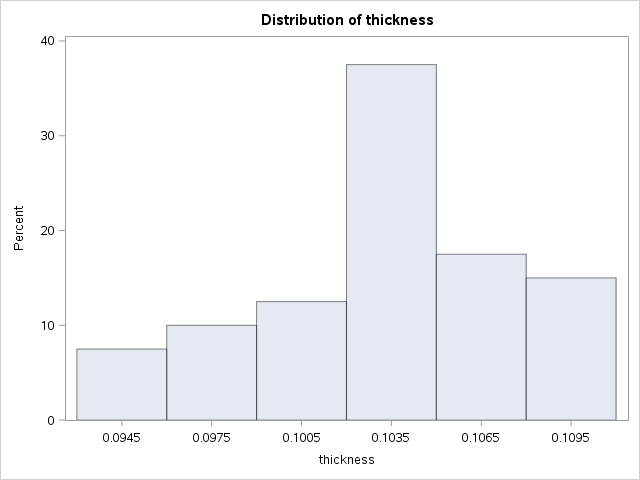
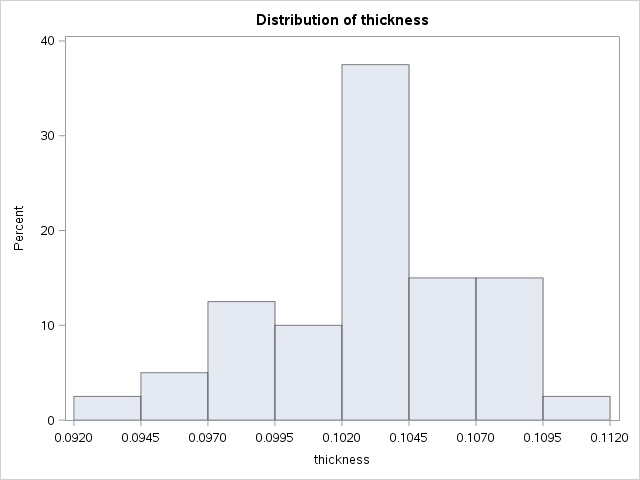
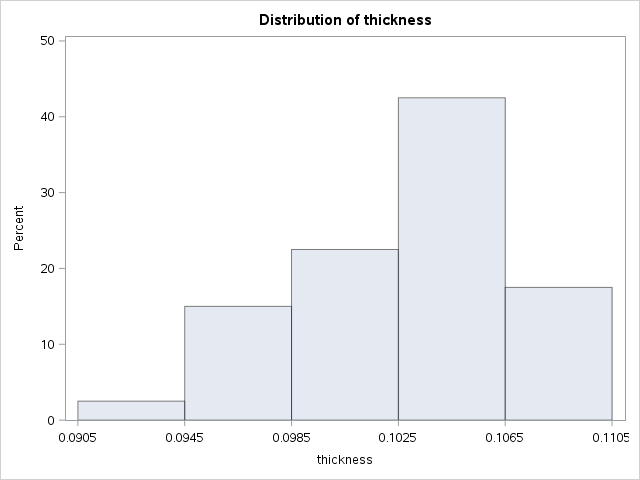
**The MEANS Procedure**

| **Variable** | **Lower 95% CL for Mean** | **Upper 95% CL for Mean** |
| --- | --- | --- |
| blue  white | 0.1038027  0.0986187 | 0.1062973  0.1023813 |

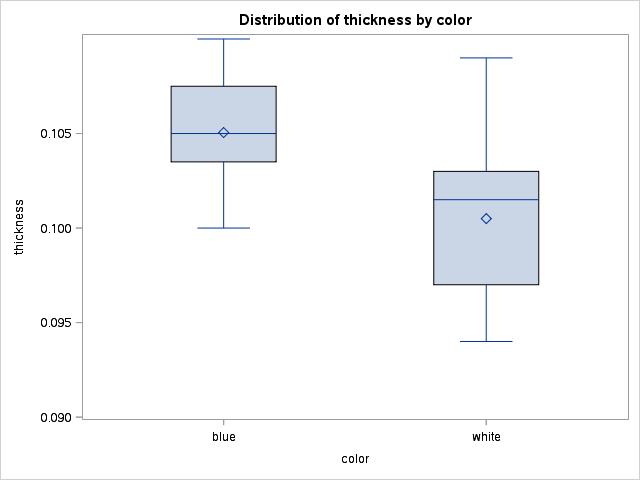
By Hand:



5:

1. 
2. 
3. 

6:



From the boxplot, I can know that there are no outliers both of the color, to blue I can know the min is 0.100, and median is 0.105, there are 10 data locate in the box, and it is skewed to left;

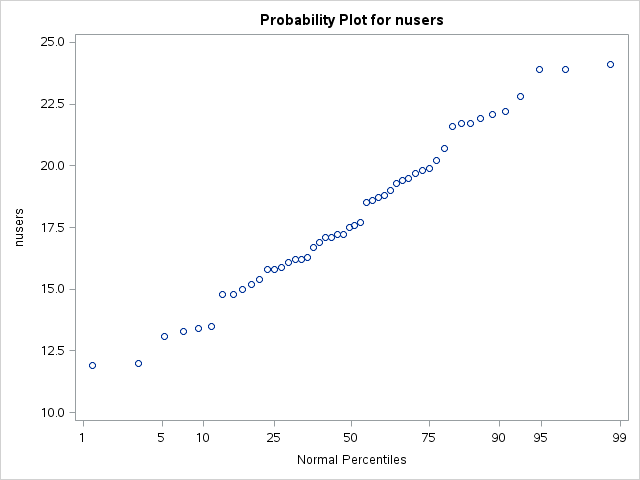
To white, I can know the min is under 0.095 and the median is larger than 0.100, it is skewed to right. According to the boxes location, I can know that, the thickness of white paper is smaller than the one of blue paper.

Part B:

1.

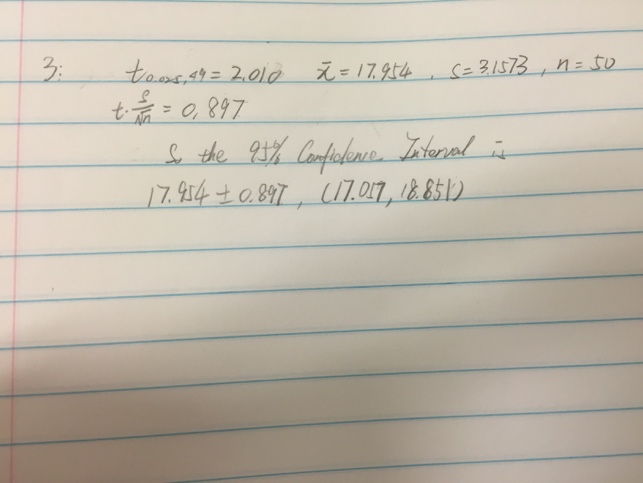
| **Obs** | **nusers** |
| --- | --- |
| **1** | 17.2 |
| **2** | 22.1 |
| **3** | 18.5 |
| **4** | 17.2 |
| **5** | 18.6 |
| **6** | 14.8 |
| **7** | 21.7 |
| **8** | 15.8 |
| **9** | 16.3 |
| **10** | 22.8 |
| **11** | 24.1 |
| **12** | 13.3 |
| **13** | 16.2 |
| **14** | 17.5 |
| **15** | 19.0 |
| **16** | 23.9 |
| **17** | 14.8 |
| **18** | 22.2 |
| **19** | 21.7 |
| **20** | 20.7 |
| **21** | 13.5 |
| **22** | 15.8 |
| **23** | 13.1 |
| **24** | 16.1 |
| **25** | 21.9 |
| **26** | 23.9 |
| **27** | 19.3 |
| **28** | 12.0 |
| **29** | 19.9 |
| **30** | 19.4 |
| **31** | 15.4 |
| **32** | 16.7 |
| **33** | 19.5 |
| **34** | 16.2 |
| **35** | 16.9 |
| **36** | 17.1 |
| **37** | 20.2 |
| **38** | 13.4 |
| **39** | 19.8 |
| **40** | 17.7 |
| **41** | 19.7 |
| **42** | 18.7 |
| **43** | 17.6 |
| **44** | 15.9 |
| **45** | 15.2 |
| **46** | 17.1 |
| **47** | 15.0 |
| **48** | 18.8 |
| **49** | 21.6 |
| **50** | 11.9 |

2.



From the graph, I can see that these spots are approaching a line, it means the data is normally distributed, and we can use t-test to know the relationship between the two variables.

3.



4.

**The TTEST Procedure**

**Variable: nusers**

| **N** | **Mean** | **Std Dev** | **Std Err** | **Minimum** | **Maximum** |
| --- | --- | --- | --- | --- | --- |
| 50 | 17.9540 | 3.1573 | 0.4465 | 11.9000 | 24.1000 |

| **Mean** | **95% CL Mean** | | **Std Dev** | **95% CL Std Dev** | |
| --- | --- | --- | --- | --- | --- |
| 17.9540 | 17.0567 | 18.8513 | 3.1573 | 2.6374 | 3.9344 |

| **DF** | **t Value** | **Pr > |t|** |
| --- | --- | --- |
| 49 | 1.69 | 0.0976 |

n = 50, x = 17.954, Sx = 3.1573, α = 0.05, SEmean = 0.447.

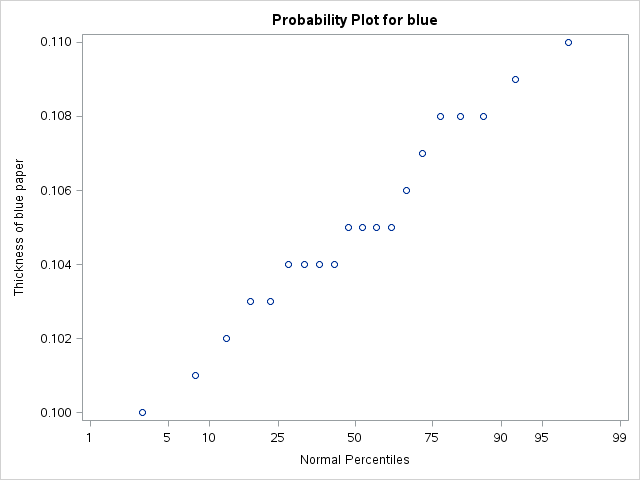
1. H0 = 17.2;

H1 != 17.2.

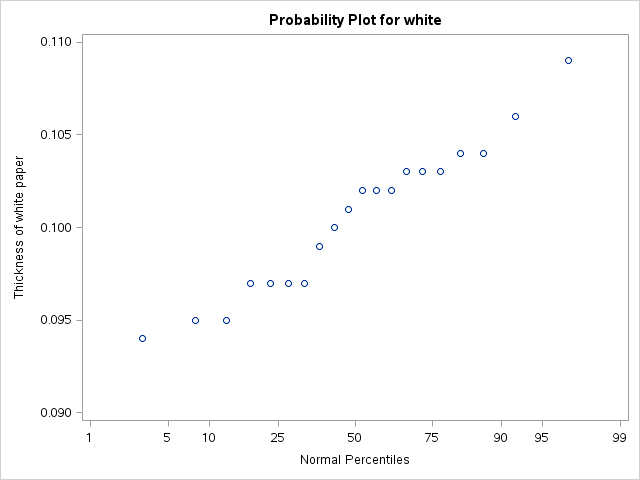
1. t = (17.954 – 17.2)/0.447 = 1.69
2. n – 1 = 49 degree of freedom, from t-table the interval is [-2.010, 2.010], it is the 95% confidence interval for t.
3. Since 1.69 ∈ [-2.010, 2.010], accept H0.
4. P = 0.0976.

Part C:

2.



The points in this graph are approaching to a line, so to blue paper, it is normally distributed.



The points in this graph are same as blue ones, approaching to a line, so to white paper, it is normally distributed.

3.

Paired-sample t-test

n = 20, d = 0.00455, Sd = 0.00353, α = 0.05, SEmean = 0.00079

1. H0: μd = 0, H1: μd != 0
2. z = 0.00455/0.00079 = 5.76
3. 95% confidence interval: n – 1 = 20 – 1 = 19 degrees of freedom, from t-table the interval is [-2.093, 2.093].
4. Since 5.76 ∉ [-2.093, 2.093], reject H0.
5. p-value < 0.0001

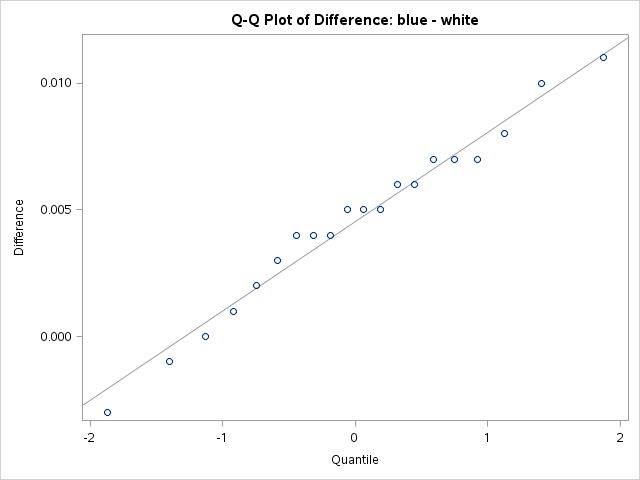
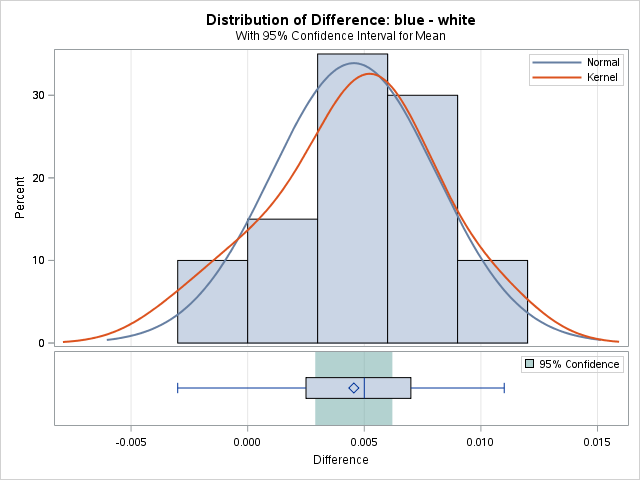
**The TTEST Procedure**

**Difference: blue - white**

| **N** | **Mean** | **Std Dev** | **Std Err** | **Minimum** | **Maximum** |
| --- | --- | --- | --- | --- | --- |
| 20 | 0.00455 | 0.00353 | 0.000790 | -0.00300 | 0.0110 |

| **Mean** | **95% CL Mean** | | **Std Dev** | **95% CL Std Dev** | |
| --- | --- | --- | --- | --- | --- |
| 0.00455 | 0.00290 | 0.00620 | 0.00353 | 0.00269 | 0.00516 |

| **DF** | **t Value** | **Pr > |t|** |
| --- | --- | --- |
| 19 | 5.76 | <.0001 |



Report:

After the t-test, H0 is rejected, so the alternative hypothesis is what I accept, which means the two different color papers’ thickness is different.

4.

Independent two-sample t-test

n1 = 20, n2 =20, x1 = 0.1051, x2 = 0.1005, S1 = 0.00267, S2 = 0.00402, α = 0.05

s = 0.00341, SEdiff = 0.00108

1. H0: μ1 = μ2, H1: μ1 != μ2
2. t = (0.1051 – 0.1005)/0.00108 = 4.26
3. n1 + n2 – 2 = 38 degrees of freedom, from t-table the interval is [-2.024, 2.024]
4. Since H0: 4.26 ∉ [-2.024, 2.024], reject H0.
5. P-value = 0.0002.

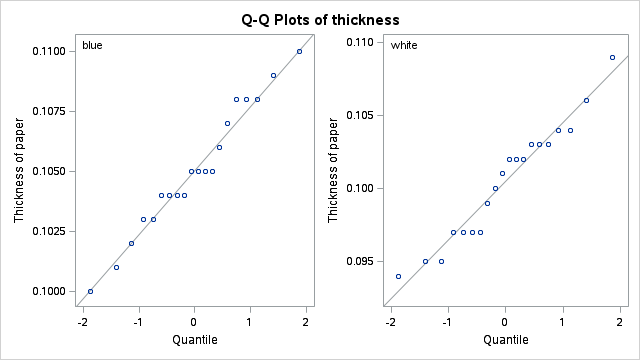
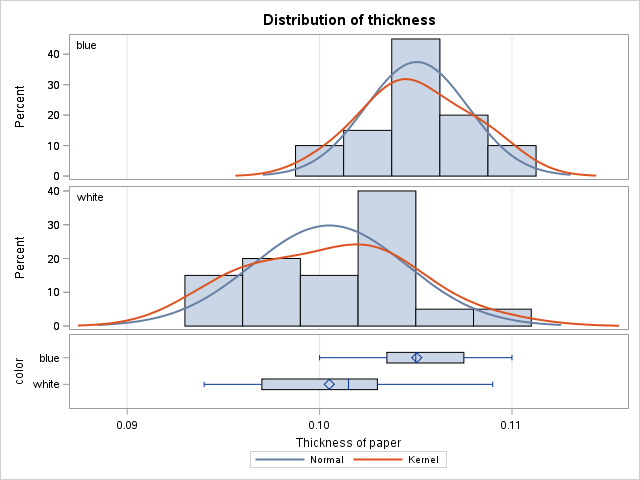
**The TTEST Procedure**

**Variable: thickness (Thickness of paper)**

| **color** | **N** | **Mean** | **Std Dev** | **Std Err** | **Minimum** | **Maximum** |
| --- | --- | --- | --- | --- | --- | --- |
| **blue** | 20 | 0.1051 | 0.00267 | 0.000596 | 0.1000 | 0.1100 |
| **white** | 20 | 0.1005 | 0.00402 | 0.000899 | 0.0940 | 0.1090 |
| **Diff (1-2)** |  | 0.00455 | 0.00341 | 0.00108 |  |  |

| **color** | **Method** | **Mean** | **95% CL Mean** | | **Std Dev** | **95% CL Std Dev** | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **blue** |  | 0.1051 | 0.1038 | 0.1063 | 0.00267 | 0.00203 | 0.00389 |
| **white** |  | 0.1005 | 0.0986 | 0.1024 | 0.00402 | 0.00306 | 0.00587 |
| **Diff (1-2)** | **Pooled** | 0.00455 | 0.00237 | 0.00673 | 0.00341 | 0.00279 | 0.00440 |
| **Diff (1-2)** | **Satterthwaite** | 0.00455 | 0.00236 | 0.00674 |  |  |  |

| **Method** | **Variances** | **DF** | **t Value** | **Pr > |t|** |
| --- | --- | --- | --- | --- |
| **Pooled** | Equal | 38 | 4.22 | 0.0001 |
| **Satterthwaite** | Unequal | 32.999 | 4.22 | 0.0002 |



After the t-test, I get that H0 is rejected, and I accept H1, which means the two different color papers’ thickness are different.

5.

After the two different t-test, I get the same result, the thickness of paper is different for the color of them. Since in both of t-test, the conclusions are all reject H1, it means the result of the tests are two kinds of paper has difference thickness. But the p-value of the two tests are different, and if p-value is much lower, the probability to reject null hypothesis could be higher. So in paired sample t-test, the p-value is much lower, it will be more likely to reject the null hypothesis; another reason, because the paired t-test use the difference between the blue and white paper, since the two papers measured by one person, if some small errors happened in the measurement, paired sample t-test will delete it from doing the reduce calculation. As a conclusion, paired sample t-test is much better in this problem, because it has a lower p-value and it will reduce the error happened in the measurement.