1. Pollution levels were measured downstream and upstream from a plant. On a given day one measure was taken upstream and one was taken downstream for in total of 16 days. Test whether there is a difference in the pollution level up and down the stream.
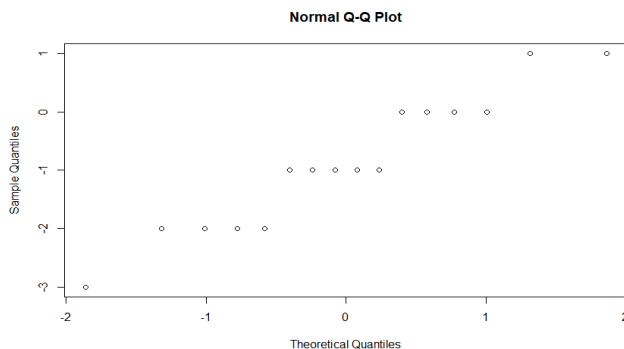
   Answer:

   Correlation of Up and Down stream measurement is 0.9861061 which means they are highly correlated. So, we do a paired test on the difference of two samples per day.

   Histogram of the difference between pollution levels of up and downstream.
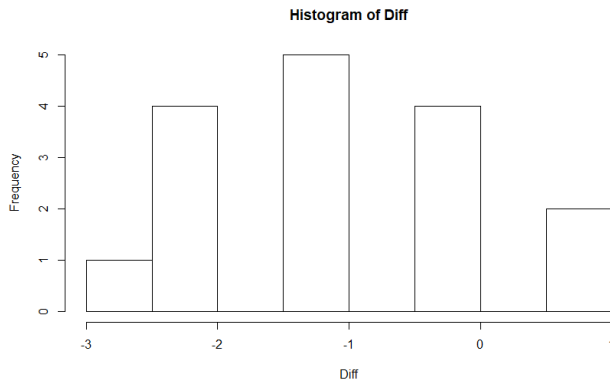
   Histogram looks symmetric.

   The Q-Q plot looks like

   Which shows a 45-degree line.

   The Shapiro-Wilk test for normality gives a p value of 0.2679 so we cannot reject the assumption of normality.

   Mean of Difference = -.875
   Variance of Difference= 1.316667
   95% confidence interval of the mean of Difference is (-1.4864388, -0.2635612)

   A paired t test on the mean of Difference between up and downstream pollution level gives a p value of 0.0081 which means we can reject the null hypothesis that the mean of Difference between up and downstream pollution are same. It is consistent with the 95% CI values.

2. Yield in pounds of tomatoes per week for two gardens with different types of fertilizer are given in the file "t.test.data.csv". Determine whether the average yields for the two gardens differ.

   Answer:

   Mean of Garden A = 3
   Variance of Garden A = 1.333

   Mean of Garden B = 5
   Variance of Garden B = 1.333

   A test for equal variances gives a p value of 1 which means we can reject the null hypothesis that the variances are equal.

   A non-parametric two sample test for equal variances gives a p value of 1 which agrees with the variance test.

   Since the variance are unequal we need to do welch test on the two means.

   With the assumption of Normality and unequal variance a unpaired t-test with the null hypothesis that the two mean are equal gives a p value of 0.001115 which means we can reject the null hypothesis that the two means are equal.

   With Normal assumption, a same test assuming equal variance gives a p value of 0.001115 which agrees with the previous test.

   Without the normal assumption a non-parametric Mann Whitney test for equal means gives a p value of 0.002988 which means we can reject the null hypothesis that the means are equal.

   So tests conclude that the average yields for the two gardens differ.