Nonparametric Statistics

Anisha Mittal

2021-02-10

```
# Loading the package for 'Probability and Statistics with R'
library(PASWR2)
```

1. Loading the dataset fertilize. (available on PASWR2 package)

```
data("FERTILIZE")
attach(FERTILIZE)
```

a. Are the samples independent or paired?

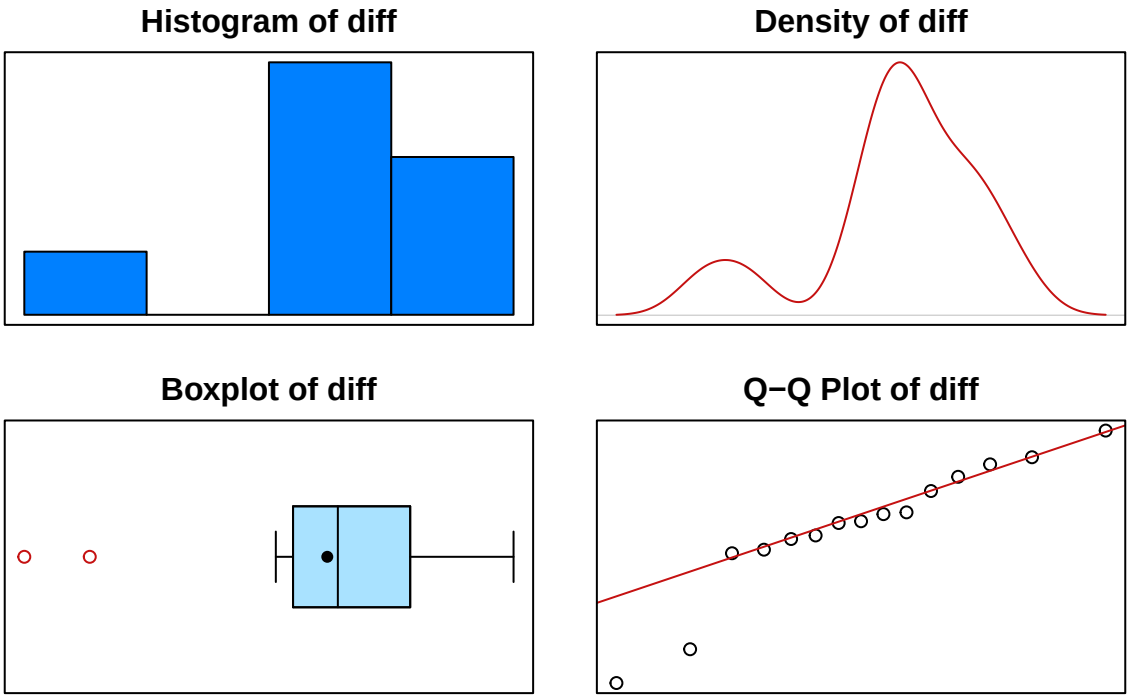
The sample is paired because even though the seeds are different they are planted in the same pot.

b. t-test to test normality

Here we are using The One Sample t-test with Null Hypothesis: mean = 0 ; Alternative Hypothesis: mean not equal to 0 ; Significance level (alpha): 0.05 ; We reject the null hypothesis if the p-value < alpha.

```
# differencing the data
diff = height[fertilization=="cross"] - height[fertilization=="self"]
eda(diff) # Exploratory Data Analysis
```

EXPLORATORY DATA ANALYSIS




```
## Size (n)  Missing  Minimum   1st Qu     Mean   Median   TrMean   3rd Qu
##   15.000    0.000   -8.375    1.375    2.617    3.000    2.617    5.625
##      Max    Stdev      Var  SE Mean   I.Q.R.    Range Kurtosis Skewness
##    9.375    4.718   22.260    1.218    4.250   17.750    0.141   -0.895
## SW p-val
##    0.098
```

```
t.test(diff , mu = 0) # Applying t-test to data
```

```
##
##  One Sample t-test
##
## data:  diff
## t = 2.148, df = 14, p-value = 0.0497
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.003899165 5.229434169
## sample estimates:
## mean of x
##  2.616667
```

Normality should not be assumed. Since, by looking at the boxplot we see the mean circle is a bit away from the median line. Also the EDA shows that mean is not equal to the median.

The One-Sample t-test gives test statistic as -2.148 the p value = 0.0497 $<$ 0.05 (alpha). Since there is sufficient evidence, we reject the null hypothesis and conclude that the alternate hypothesis is probably true.

c. Wilcoxone Signed-rank test

Here we are using The Wilcoxon signed-rank test with Null Hypothesis: Mean = 0 ; Alternative Hypothesis: Mean not equal to 0 ; Significance level (alpha): 0.05 ; We reject the null hypothesis if the p-value < alpha

```
# Using wilcoxone signed-rank test and removing ties
wilcoxe.test(diff, mu= 0)
```

```
##
##  Wilcoxon Signed Rank Test
##
## data:  diff
## t+ = 96, p-value = 0.04126
## alternative hypothesis: true median is not equal to 0
## 95.20874 percent confidence interval:
##  0.5000 5.1875
## sample estimates:
## (pseudo)median
##          3.125
```

The Wilcoxon signed-rank test gives test statistic as 96 and the p value = 0.04126 < 0.05 (alpha). Since there is sufficient evidence, we reject the null hypothesis and conclude that the alternate hypothesis is probably true.

d. Permutation test

Here we are using The Permutation test with Null Hypothesis: Mean = 0 ; Alternative Hypothesis: Mean not equal to 0 ; Significance level (alpha): 0.05 ; We reject the null hypothesis if the p-value < alpha

```
# Creating a funtion for the test
binary<-function(y,digits)
  {
  ans<-0:(digits-1)
  (y %/% 2^ans)%%2
  }


digits<-length(diff)
n<-2^digits
perm.res=numeric(n) #Create a vector of length n:

for (i in 1:n){
  x<-diff*2*(binary(i,digits)-0.5)
  perm.res[i]<-mean(x)
}

observed=mean(diff)

observed
```

```
## [1] 2.616667
```

```r
pvalue=mean(perm.res >= observed)
pvalue
```

```
## [1] 0.02633667
```

The permutation test gives p value = 0.02633667 < 0.05 (alpha). Since there is sufficient evidence, we reject the null hypothesis and conclude that the alternate hypothesis is probably true.

e. Are all the p-values for the above tests performed same?

No, we dont get the same p-values from all the tests. We get a p-value of 0.0497 when one-sample t-test is applied, a p-value of 0.04126 when the wilcoxone signed-rank test is applied, and a p-value of 0.02633667 when a permutation test is applied. This suggests that the permutation test gives more accurate results and this test should be considered.