Alla Topp

MSDS 660, Nonparametric statistics

1) **You want to analyze a new chain restaurant "Souperb". 15 people are asked to rate the food on a 5-point scale (1=terrible, 5=excellent). The ratings are 5,3,2,1,4,3,5,1,5,2,3,4,2,1,3. Use the Sign Test to test the median rating whether it is at least 3. Given 95% confidence interval. State your hypothesis, test statistics, p-value, and conclusion.**

Hypothesis:

H0: the median < 3.

H1: the median >= 3.

```
> a <- c(5,3,2,1,4,3,5,1,5,2,3,4,2,1,3) # numeric vector of data values.
> SIGN.test(a, mu = 3, conf.level = 0.95, alternative = "greater") # one sample sign test

        One-sample Sign-Test

data:  a
s = 15, p-value = 3.052e-05
alternative hypothesis: true median is greater than 0
95 percent confidence interval:
   2 Inf
sample estimates:
median of x
          3

Achieved and Interpolated Confidence Intervals:

                Conf.Level L.E.pt U.E.pt
Lower Achieved CI   0.9408      2    Inf
Interpolated CI     0.9500      2    Inf
Upper Achieved CI   0.9824      2    Inf
```

Conclusion:

For this problem I used one sample sign.test() (we have only one sample of ratings). We have 15 ratings and the median between positive and negative ratings is 3 (that`s why we have mu=3). I stated that alternative is "greater" because my alternative hypothesis is if the median is greater or equal than 3. Looking at the results we can see that p-value turns out to be less than 0.05 significance level. We reject the null hypothesis and accept alternative hypothesis that the ratings are larger or equal than 3.

2) **2) Two different operating systems (M, W) are rated using 1-10 scale (the higher the better rate). The rates of M operating system are: 9,8,5,3,6,10,4,2,8,7. The rates for W operating system are: 7,6,8,2,9,5,1,4,7,10. Do these operating systems have the same distribution? What statistical test will you use? Given the significance level = 0.05. State your hypothesis, test statistics, p-value, and conclusion.**

Hypothesis:

H0: There is no difference between two operating system's ratings.

H1: There is a difference between two operating system's ratings.

Alla Topp
MSDS 660, Nonparametric statistics

 To test the hypothesis, we apply the wilcox.test function to compare the independent samples.

Distribution

Yes, two operating systems have the same normal distribution.

What statistical test?

Mann-Whitney-Wilcoxon Test or Mann-Whitney U-Test (two independent samples)

Conclusion:

```
> #problem 2
> m <- c(9,8,5,3,6,10,4,2,8,7) # numeric vector with rates of M operating system
> w <- c(7,6,8,2,9,5,1,4,7,10) # rates of W operating system
> wilcox.test(m, w, mu = 0, paired=TRUE, exact = FALSE, conf.int=TRUE) # computes wilcox test

        Wilcoxon signed rank test with continuity correction

data:  m and w
V = 28.5, p-value = 0.959
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -2.499951  2.500014
sample estimates:
(pseudo)median
  7.914168e-06
```

My p-value is larger than 0.05, it means that there is no significant difference between ratings of two operating systems. The results state that we accept the null hypothesis. On the other hand, looking at the rating, I have a feeling that two operating systems should not be identical. I tried to calculate the difference between two operating systems myself and saw the difference in means. This is my concern for this problem.