1. The following data gives the stroke index for 10 patients before and after treatment. We wish to test the hypothesis that the treatment has no effect.

| Before | 109 57 53 57 68 72 51 65 52 61 |
|        | After | 56 44 55 40 62 46 48 41 56 49 |

(a) Calculate the differences, draw a boxplot of these and comment on why a non-parametric test might be preferred.

(b) Use the sign test to see if the treatment has no effect.

(c) Perform a Wilcoxon sign-rank test to test the same hypothesis as in (b).

Solution:

(a)

<table>
<thead>
<tr>
<th>Difference($x_i - y_i$)</th>
<th>53 13 -2 17 6 26 3 24 -4 12</th>
</tr>
</thead>
</table>

Comment: The boxplot looks not symmetric. Hence it is doubtful to assume the sample is from a normal population. On the other hand, the sample size is small. So it might be better to perform a nonparametric test.

(b) The hypothesis to be tested is

$$H : \mu = 0 \quad \text{vs} \quad H_A : \mu \neq 0,$$

where $\mu$ is the median of the difference between before and after treatment.

Let $X$ be the number of the positive $x_i - y_i$. Then a observed value of the $X$ is 8, and hence the $p$-value of the sign test is

$$p = 2P(B(10, 1/2) \geq 8) = 2P(B(10, 1/2) \leq 2) = 2 \times 0.0547 = 0.1094.$$  

The $p$-value is greater than 0.10. Hence the data is consistent with the null hypothesis $H$.

(c) **Step 1.** Find the ranks of the $|x_i - y_i|$.

<table>
<thead>
<tr>
<th>Difference($x_i - y_i$)</th>
<th>53 13 -2 17 6 26 3 24 -4 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference($</td>
<td>x_i - y_i</td>
</tr>
<tr>
<td>Ranks</td>
<td>10 6 1 7 4 9 2 8 3 5.</td>
</tr>
</tbody>
</table>
Step 2. Calculate $EW^+$ and $\text{var}(W^+)$ (Note that there are no ties on the data)

$$
EW^+ = 10 \times (10 + 1)/4 = 110/4, \\
\text{var}(W^+) = 10 \times (10 + 1) \times (2 \times 10 + 1)/24 = 385/4.
$$

Step 3. Calculate $w^-, w^+$ and $w$.

$$
w^- = 1 + 3 = 4, \\
w^+ = 10 \times (10 + 1)/2 - w^- = 51, \\
w = \min\{w^-, w^+\} = 4.
$$

Step 4. Calculate the $p$-value:

$$
p = 2\text{P}(W^+ \leq 4) = 2 \times 0.0068 = 0.0136.
$$

Conclusion: There is strong evidence to reject the null hypothesis $H$ based on the Wilcoxon sign-rank test.

2. When there are no ties on the data, the Wilcoxon sign-rank test statistic may be written as

$$
W^+ = I_1 + 2I_2 + \cdots + nI_n,
$$

where $I_1, \ldots, I_n$ are iid $B(1,1/2)$ rv’s. Show that

$$
EW^+ = \frac{1}{2}(1 + \cdots + n) = \frac{n(n + 1)}{4}, \\
\text{var}(W^+) = \frac{1}{4}(1^2 + \cdots + n^2) = \frac{n(n + 1)(2n + 1)}{24}.
$$

Solution: Note that $I_k \sim B(1,1/2)$. We get for any $1 \leq k \leq n$

$$
E(I_k) = 1/2, \quad \text{var}(I_k) = 1/4.
$$

This yields that

$$
EW^+ = E(I_1) + 2E(I_2) + \cdots + nE(I_n) = \frac{1}{2}(1 + 2 + \cdots + n) = \frac{n(n + 1)}{4}, \\
\text{var}(W^+) = \text{var}(I_1) + 2^2\text{var}(I_2) + \cdots + n^2\text{var}(I_n) = \frac{1}{4}(1^2 + \cdots + n^2) = \frac{n(n + 1)(2n + 1)}{24}.
$$
Computer Exercises solution week 3

1. The data in \(amp\) are the results of measurements made using two methods on 15 pairs of tablets to determine the dosage of ampicillin. They were analysed in week 1 using a t-test.

(a) Use the sign test to see if there is a systematic difference between the two measurement methods. Hint: You may use the following steps in Splus.

i. State the hypothesis to be tested
ii. Create a vector \(d\) whose elements correspond to the difference of two columns of the data \(amp\). (Note that there is a zero observation in \(d\))
iii. Remove zero from \(d\), and name the new vector as \(d1\).
iv. Calculate the size \(n\) of \(d1\) and the number \(x\) of positive elements in \(d1\).
v. Calculate the \(p\)-value by using \(pbinom\) or \(binom.test\).

(b) Use the Wilcoxon sign-rank test to test the same hypothesis as in (a).

i. Find the ranks of the \(|d|\). [Recall \(d\) is defined as in part (a)] Name the rank vector as \(r\).
ii. Calculate the \(EW^+\) and \(var(W^+)\).
iii. Calculate the \(w^-\) and \(w^+\).
iv. Calculate the \(w\) and then the \(p\)-value.
(Note that there is a zero observation in \(d\))

(c) Use \(wilcox.test\) to confirm your findings in (b).
(Note that we didn’t consider any corrections in part (b))

(d) Perform a t-test for the same hypothesis as in (a).

(e) Compare the results obtained by three different tests: the sign test, the Wilcoxon sign-rank test and the t-test. Comment on the differences in the three different tests including the assumptions and the \(p\)-values.

(f) Obtain a boxplot of \(d\). Is a transformation necessary to test the hypothesis in (a) by using the Wilcoxon signed-rank test or the t-test?

Solution: In Splus

#(a)

(i) \(H: \mu=0 \ vs \ H_A: \mu\not=0\)
where \(\mu\) denotes the mean difference of two measurements.

(ii)
\[d<-amp[,1]-amp[,2]\]

(iii)
d1<-d[d!=0]

(iv)
  n<-length(d1)
  x<-sum(d1>0) or x<-length(d1[d1>0])

(v)
  p<-1  # since x=7=n/2

  or
  > binom.test(x, n, p=0.5, alter="two.sided")

  Exact binomial test
data:  x out of n
  number of successes = 7, n = 14, p-value = 1
  alternative hypothesis: true p is not equal to 0.5

#(b)
(i)
r<-rank(abs(d))

(ii)
  ew<-sum(r[d!=0])/2
  vw<-sum((r[d!=0])^2)/4

(iii)
  w.plus<-sum(r[d>0])
  w.minus<-sum(r[d<0])

(iv)
  w<-min(w.plus, w.minus)
  p<-2*pnorm((w-ew)/sqrt(vw))
  p
  [1] 0.6290533

#c)
  wilcox.test(d, alternative="two.sided",
               mu=0, exac=F, correct=F)

  4
Wilcoxon signed-rank test

data:  d
signed-rank normal statistic without correction
Z = 0.4831, p-value = 0.6291
alternative hypothesis: true mu is not equal to 0

or
wilcox.test(amp[,1], amp[,2], alternative=
  "two.sided", mu=0, paired=T, exact=F, correct=T)

#(d)
t.test(amp[,1], amp[,2], alternative="two.sided",
  mu=0, paired=T)

or
t.test(d, alternative="two.sided", mu=0)

One-sample t-Test

data:  d
t = 0.368, df = 14, p-value = 0.7184
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -2.124433  3.004433
sample estimates:
 mean of x
  0.44

#(e) All three tests give the same conclusion that the data are consistent with the null hypothesis in (a). However the p-values are quite different. The reasons behind the differences of the p-values are mainly because the assumptions used for the three different tests are quite different. For the t-test we assume the sample is from a normal population. For the Wilcoxon sign rank test, we only assume the sample is from a symmetric population. Note that the assumption required for the sign test is weakest. However the test might lack power.

#(f)
boxplot(d)
The boxplot looks quite symmetric. Transformation is not necessary before using the Wilcoxon signed-rank test and the $t$-test.