2.3. Nonparametric methods

1. One-sample problem:

$X_1, X_2, \cdots, X_n$ are random variables with common distribution (not necessarily normal distribution). Suppose M is the median of the distribution. We want to test if $H_0: M = M_0$, where M_0 is some number. We now introduce the Wilcoxon sign rank test.

Wilcoxon sign rank test:

Let $Z_i = |X_i - M_0|$ and R_i is the rank of Z_i. Let $T^+ = \sum_{(X_i - M_0)>0} R_i$ be the sum of all the ranks corresponding to $X_i - M_0 > 0$, while $T^- = \sum_{(X_i - M_0)<0} R_i$ is the sum of all the ranks corresponding to $X_i - M_0 < 0$.

Example :

$H_0: M = M_0 = 8$ and the following table:

X_i	7.1	13.5	5.2	4.2	6.4	10.4	5.7	5.4	6.3	7.5
Z_i	0.9	5.5	2.8	3.8	1.6	2.4	2.3	2.6	1.7	0.5
R_i	2	10	8	9	3	6	5	7	4	1
Sign of $X_i - 8$	$-$	$+$	$-$	$-$	$-$	$+$	$-$	$-$	$-$	$-$

Then,

$$T^+ = 10 + 6 = 16, T^- = 2 + 8 + 9 + 3 + 5 + 7 + 4 + 1 = 39.$$

The following are the Wilcoxon sign rank tests under three alternatives:

1. $H_0: M = M_0 \ vs. H_1: M < M_0 \Rightarrow T^+ \leq W_{n,\alpha}$ then reject H_0, where $W_{n,\alpha}$ can be found by table.
2. $H_0: M = M_0 \ vs. H_1: M > M_0 \Rightarrow T^- \leq W_{n,\alpha}$ then reject H_0.
3. $H_0: M = M_0 \ vs. H_1: M \neq M_0 \Rightarrow min(T^+, T^-) \leq W_{n,\alpha}$ then reject H_0.

Note:

$$T^+ + T^- = \frac{n(n+1)}{2}.$$

Example (continue):

Test $H_0: M = M_0 = 8 \ vs. H_1: M < 8, \ \alpha = 0.05.$

[solution]:

$$T^+ = 16, W_{10,0.05} = 11 \Longrightarrow T^+ > W_{10,0.05},$$

we do not reject H_0.

Example (Splus):

```
s=c(7.1,13.5,5.2,4.2,6.4,10.4,5.7,5.4,6.3,7.5)
wilcox.test(s,mu=8,alt="less")
help(wilcox.test)
```

2. Two-samples problem:

$X_1, X_2, \cdots, X_n \sim F_1$ and $Y_1, Y_2, \cdots, Y_m \sim F_2$, where F_1 and F_2 are some distributions. Let M_1 and M_2 be the medians of F_1 and F_2, respectively. We want to test $H_0: M_1 = M_2$.

Wilcoxon rank sum test:

Step 1: Combine $X_1, X_2, \cdots, X_n, Y_1, Y_2, \cdots, Y_m$ and find all the ranks of these data.

Step 2: Find the sum of the ranks W corresponding to $X_1, X_2, \cdots, X_n$.

Step 3: The following are the Wilcoxon rank sum tests under 3 alternatives:

1. $H_0: M_1 = M_2$ $vs. H_1: M_1 > M_2 \Rightarrow W > W_U$ then reject H_0, where W_U can be found by table.
2. $H_0: M_1 = M_2$ $vs. H_1: M_1 < M_2 \Rightarrow W < W_L$ then reject H_0, where W_L can be found by table.
3. $H_0: M_1 = M_2$ $vs. H_1: M_1 \neq M_2 \Rightarrow \mathrm{W} > W_U$ or $W < W_L$, then reject H_0.

Example:

South	60	100	150	290		
North	110	130	140	170	200	310

The above data are the incomes in both North and South areas. Test

$$H_0: M_1 = M_2 \ \ vs. H_1: M_1 \neq M_2$$

with $\alpha = 0.05$.

[solution]:

Step 1:

Income	60	100	150	290	110	130	140	170	200	310
Rank	1	2	6	9	3	4	5	7	8	10

Step 2:

$$\mathrm{n} = 4, m = 6, W = 1 + 2 + 6 + 9 = 18, W_L = 12, W_U = 32$$

Step 3:

$$12 = W_L \leq W = 18 \leq 32 = W_U,$$

we do not reject H_0

Example (Splus):

```
in1=c(60,100,150,290)
in2=c(110,130,140,170,200,310)
wilcox.test(in1,in2)
```

Useful links:

Core Ftp: http://coreftp.com/download.html

Recuva: http://www.piriform.com/recuva/download