

1、 Use t test, Wilcoxon signed rank test, sign test to validate results of the 95% and 99% confidence intervals on ANSM p. 75.

Code :

```
1 library(DescTools)
2 x<-c(-2 ,4 ,8 ,25 ,-5 ,16 ,3 ,1 ,12 ,17 ,20 ,9)
3 x1<-c(-2 ,4 ,8 ,35 ,-5 ,16 ,3 ,1 ,12 ,17 ,20 ,9)
4 x2<-c(-2 ,4 ,8 ,65 ,-5 ,16 ,3 ,1 ,12 ,17 ,20 ,9)
5 t.test(x1)
6 t.test(x1,conf.level = 0.99)
7 wilcox.test(x1,conf.int = TRUE)
8 wilcox.test(x1,conf.int = TRUE,conf.level = 0.99)
9 SignTest(x1)
10 SignTest(x1,conf.level = 0.99)
11
12 t.test(x2)
13 t.test(x2,conf.level = 0.99)
14 wilcox.test(x2,conf.int = TRUE)
15 wilcox.test(x1,conf.int = TRUE,conf.level = 0.99)
16 SignTest(x2)
17 SignTest(x2,conf.level = 0.99)
```

Result :

```
1 > t.test(x1)
2
3 One Sample t-test
4
5 data: x1
6 t = 3.0707, df = 11, p-value = 0.01065
7 alternative hypothesis: true mean is not equal to 0
8 95 percent confidence interval:
9 2.785007 16.881660
10 sample estimates:
11 mean of x
12 9.833333
13
14 > t.test(x1,conf.level = 0.99)
15
16 One Sample t-test
17
18 data: x1
```

```

19 t = 3.0707, df = 11, p-value = 0.01065
20 alternative hypothesis: true mean is not equal to 0
21 99 percent confidence interval:
22 -0.1125486 19.7792152
23 sample estimates:
24 mean of x
25 9.833333
26
27 > wilcox.test(x1,conf.int = TRUE)
28
29 Wilcoxon signed rank test
30
31 data: x1
32 V = 71, p-value = 0.009277
33 alternative hypothesis: true location is not equal to 0
34 95 percent confidence interval:
35 2.5 17.0
36 sample estimates:
37 (pseudo)median
38 9
39
40 > wilcox.test(x1,conf.int = TRUE,conf.level = 0.99)
41
42 Wilcoxon signed rank test
43
44 data: x1
45 V = 71, p-value = 0.009277
46 alternative hypothesis: true location is not equal to 0
47 99 percent confidence interval:
48 0.5 20.0
49 sample estimates:
50 (pseudo)median
51 9
52
53 > SignTest(x1)
54
55 One-sample Sign-Test
56
57 data: x1
58 S = 10, number of differences = 12, p-value = 0.03857
59 alternative hypothesis: true median is not equal to 0
60 96.1 percent confidence interval:
61 1 17
62 sample estimates:
63 median of the differences
64 8.5
65
66 > SignTest(x1,conf.level = 0.99)
67

```

```

68     One-sample Sign-Test
69
70 data:  x1
71 S = 10, number of differences = 12, p-value = 0.03857
72 alternative hypothesis: true median is not equal to 0
73 99.4 percent confidence interval:
74   -2 20
75 sample estimates:
76 median of the differences
77           8.5
78
79 >
80 > t.test(x2)
81
82     One Sample t-test
83
84 data:  x2
85 t = 2.3331, df = 11, p-value = 0.03965
86 alternative hypothesis: true mean is not equal to 0
87 95 percent confidence interval:
88   0.6983744 23.9682923
89 sample estimates:
90 mean of x
91   12.33333
92
93 > t.test(x2,conf.level = 0.99)
94
95     One Sample t-test
96
97 data:  x2
98 t = 2.3331, df = 11, p-value = 0.03965
99 alternative hypothesis: true mean is not equal to 0
100 99 percent confidence interval:
101  -4.084738 28.751405
102 sample estimates:
103 mean of x
104   12.33333
105
106 > wilcox.test(x2,conf.int = TRUE)
107
108     Wilcoxon signed rank test
109
110 data:  x2
111 V = 71, p-value = 0.009277
112 alternative hypothesis: true location is not equal to 0
113 95 percent confidence interval:
114   2.5 18.5
115 sample estimates:
116 (pseudo)median

```

```

117          9
118
119 > wilcox.test(x1,conf.int = TRUE,conf.level = 0.99)
120
121     Wilcoxon signed rank test
122
123 data:  x1
124 V = 71, p-value = 0.009277
125 alternative hypothesis: true location is not equal to 0
126 99 percent confidence interval:
127    0.5 20.0
128 sample estimates:
129 (pseudo)median
130          9
131
132 > SignTest(x2)
133
134     One-sample Sign-Test
135
136 data:  x2
137 S = 10, number of differences = 12, p-value = 0.03857
138 alternative hypothesis: true median is not equal to 0
139 96.1 percent confidence interval:
140    1 17
141 sample estimates:
142 median of the differences
143          8.5
144
145 > SignTest(x2,conf.level = 0.99)
146
147     One-sample Sign-Test
148
149 data:  x2
150 S = 10, number of differences = 12, p-value = 0.03857
151 alternative hypothesis: true median is not equal to 0
152 99.4 percent confidence interval:
153   -2 20
154 sample estimates:
155 median of the differences
156          8.5

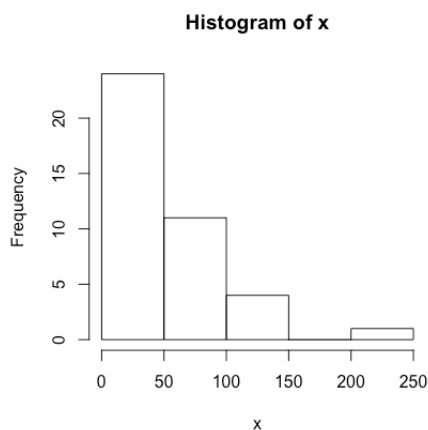
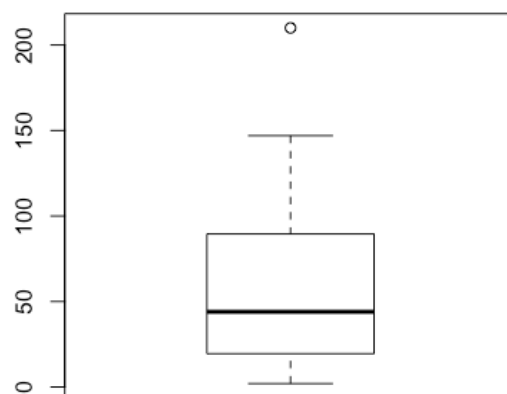
```

2、ANSM 3.16

A pathologist counts the numbers of diseased plants in randomly selected areas each 1 metre square on a large field. For 40 such areas the numbers of diseased plants are:

21	18	42	29	81	12	94	117	88	210
44	39	11	83	42	94	2	11	33	91
141	48	12	50	61	35	111	73	5	44
6	11	35	91	147	83	91	48	22	17

Use histograms and boxplots to decide whether there is evidence of skewness or outliers. Use nonparametric tests to find whether it is reasonable to assume the median number of diseased plants per square metre might be 50 (i) without assuming population symmetry, (ii) assuming population symmetry. For these data and the evidence provided by a boxplot and histograms do you consider the latter assumption reasonable?



first there exist an outlier and it is reasonable to assume the median number of diseased plants per square metre might be 50. secondly, assuming population symmetry is unreasonable.

Code :

```

1 x<-c(21 , 18 , 42 , 29 , 81 , 12 , 94 , 117 , 88 , 210 ,
2      44 , 39 , 11 , 83 , 42 , 94 , 2 , 11 , 33 , 91 ,
3      141 , 48 , 12 , 50 , 61 , 35 , 111 , 73 , 5 , 44 ,
4      6 , 11 , 35 , 91 , 147 , 83 , 91 , 48 , 22 , 17)
5 boxplot(x)
6 hist(x)

```

3、ANSM 4.8

The journal Biometrics 1985, 41, p. 830, gives data on numbers of medical papers published annually in that journal for the period 1971 – 81. These data are extended below to cover (in order) the period 1969 – 85. Is there evidence of a monotonic trend in numbers of medical papers published?

11 6 14 13 18 14 11 22 19 19 25 24 38 19 25 31 19

```

1 cox.stuart.test = function(x) {
2   method = "Cox-Stuart test for trend analysis"
3   leng = length(x)
4   apross = round(leng)%2
5   if (apross == 1) {
6     delete = (length(x) + 1)/2
7     x = x[-delete]
8   }
9   half = length(x)/2
10  x1 = x[1:half]
11  x2 = x[(half + 1):(length(x))]
12  difference = x1 - x2
13  signs = sign(difference)
14  signcorr = signs[signs != 0]
15  pos = signs[signs > 0]
16  neg = signs[signs < 0]
17  if (length(pos) < length(neg)) {
18    prop = pbinom(length(pos), length(signcorr), 0.5)
19    names(prop) = "Increasing trend, p-value"
20    rval <- list(method = method, statistic = prop)
21    class(rval) = "htest"
22    return(rval)
23  } else {
24    prop = pbinom(length(neg), length(signcorr), 0.5)
25    names(prop) = "Decreasing trend, p-value"
26    rval <- list(method = method, statistic = prop)
27    class(rval) = "htest"
28    return(rval)
29  }
30 }

```

```

1 > x<-c(11 , 6 , 14 , 13 , 18 , 14 , 11 , 22 , 19 , 19 , 25 , 24 , 38 , 19 ,
2 25 , 31 , 19)
3
4 Cox-Stuart test for trend analysis
5
6 data:
7 Increasing trend, p-value = 0.035156
8

```

We think the numbers of medical papers published is in Increasing trend.

4、 ANSM 4.11

A psychologist is testing 16 applicants for a job one at a time. Each has to perform a series of tests and the psychologist awards an overall point score to each applicant. A high score indicates a good performance. As each applicant may discuss the tests with later applicants before the latter are tested it is suggested that those tested later may have an unfair advantage. Do the applicants' scores (in order of testing) given below support this assertion?

62 69 55 71 64 68 72 75 49 74 81 83 77 79 89 42

Use an appropriate runs test. Do you consider the Cox-Stuart test (Section 3.2.3) may also be appropriate? Give reasons for your decision.

Code :

```

1 x<-c(62 , 69 , 55 , 71 , 64 , 68 , 72 , 75 , 49 , 74 , 81 , 83 , 77 , 79 ,
2 89 , 42)
3 cox.stuart.test(x)
4 l=length(x)
5 xs=sort(x)
6 x1=c()
7 x2=c()
8 t1=1
9 t2=1
10 for( i in 1:16){
11   if(x[i]<=71){
12     x1[t1]=x[i]
13     t1=t1+1;
14   }else{
15     x2[t2]=x[i]
16     t2=t2+1;
17   }
18 }
19 cox.stuart.test(x1)
20 cox.stuart.test(x2)

```

Result :

```
1 > cox.stuart.test(x)
2
3   Cox-Stuart test for trend analysis
4
5 data:
6 Increasing trend, p-value = 0.14453
7
8 > cox.stuart.test(x2)
9
10  Cox-Stuart test for trend analysis
11
12 data:
13 Increasing trend, p-value = 0.0625
14
15 > cox.stuart.test(x1)
16
17  Cox-Stuart test for trend analysis
18
19 data:
20 Decreasing trend, p-value = 0.3125
21
```

We divide the data into two parts and do Cox-Stuart test separately. We found that the results are quite different, Cox-Stuart test may be less powerful as monotonicity assumption is unlikely to be justified.

5、ANSM 5.6

One hundred general practitioners attend a health promotion workshop. At the start of the workshop they are asked to indicate whether they are in favour of routinely asking patients about alcohol consumption. They are then shown a video on the health and social problems caused by the excessive consumption of alcoholic drinks. The video is followed by discussion in small groups. After the video and discussion they are asked the original question again. Do the results given below indicate a significant change in attitudes as a result of the video and group discussion?

Before video and discussion			
		In favour	Against
After video and discussion	In favour	41	27
	Against	16	58

```
1 > x<-matrix(c(41,27,16,58),2,2)
2 > mcnemar.test(x)
3
4 McNemar's Chi-squared test with continuity
5 correction
6
7 data: x
8 McNemar's chi-squared = 2.3256, df = 1, p-value= 0.1273
```

No strong evidence of change in attitudes.