STA 104 Applied Nonparametric Statistics

Chapter 2: One-Sample Methods for Location Problem

Xiner Zhou

Department of Statistics, University of California, Davis

Table of contents

Paired Comparisons

Setting

We obtain $2n$ observations, two observations on each of n subjects (blocks, patients, etc.).

Subject i	X_i	Y_i
1	X_1	Y_1
2	X_2	Y_2
.	.	.
.	.	.
.	.	.
n	X_n	Y_n

- effect of medical treatment
- effect of new product
- effect of media on public opinion

- The differences $Z_1 = Y_1 - X_1, \ldots, Z_n = Y_n - X_n$ are mutually independent.
- $Z, i = 1, \ldots, n$, comes from a continuous population with a median θ, treatment effect.

Hypothesis

The null hypothesis of interest here is that of zero shift in location due to the treatment, namely,

$$H_0 : \theta = 0.$$

Two-Sided Test:

$$H_0 : \theta = 0 \text{ versus } H_a : \theta \neq 0$$

One-Sided Upper-Tail Test:

$$H_0 : \theta = 0 \text{ versus } H_a : \theta > 0$$

One-Sided Lower-Tail Test:

$$H_0 : \theta = 0 \text{ versus } H_a : \theta < 0$$

Procedure

One-sample tests for location!

Wilcoxon Signed Rank Test, Fisher's Signed Test

Example: Hamilton Depression Scale Factor

The data are a portion of the data obtained by Salsburg (1970). These data, based on nine patients who received tranquilizer, were taken from a double-blind clinical trial involving two tranquilizers. The measure used was the Hamilton depression scale factor (the "suicidal" factor). The X (pre) value was obtained at the first patient visit after initiation of therapy, whereas the Y (post) value was obtained at the second visit after initiation of therapy. The patients had been diagnosed as having mixed anxiety and depression.

The question of interest is whether the tranquilizer reduce Hamilton depression scale.

$$H_0 : \theta = 0, H_a : \theta < 0$$

Patient i	X_i	Y_i
1	1.83	0.878
2	0.50	0.647
3	1.62	0.598
4	2.48	2.05
5	1.68	1.06
6	1.88	1.29
7	1.55	1.06
8	3.06	3.14
9	1.30	1.29

Wilcoxon's signed rank test

Exact test:

```
> pre<-c(1.83, .50,1.62,2.48,1.68,1.88,1.55,3.06,1.30)
> post<-c(.878, .647, .598,2.05,1.06,1.29,1.06,3.14,1.29)
> z=post-pre
> z
[1] -0.952  0.147 -1.022 -0.430 -0.620 -0.590 -0.490  0.080 -0.010
> sort(abs(z))
[1] 0.010 0.080 0.147 0.430 0.490 0.590 0.620 0.952 1.022
```

$$T^+ = 5$$

$$p - value = 10/2^9 = 0.01953125$$

$T^+ =$	# configurations
0	(.)
1	(1)
2	(2)
3	(3), (1,2)
4	(4), (1,3)
5	(5), (1,4) (2,3)

Confirm with built-in function:

```
> wilcox.test(post,pre,paired = T,alternative = "less")

        Wilcoxon signed rank exact test

data:  post and pre
V = 5, p-value = 0.01953
alternative hypothesis: true location shift is less than 0
```

Wilcoxon's signed rank test

Large-sample approximation:

$$T^* = \frac{T^+ - \frac{n(n+1)}{4}}{\left\{\frac{n(n+1)(2n+1)}{24}\right\}^{1/2}} = \frac{5 - (9(10)/4)}{\{9(10)(19)/24\}^{1/2}} = -2.07$$

$$p - value = P(Z < -2.07) = 0.01922617$$

```
> pnorm(-2.07)
[1] 0.01922617
```

Both the exact test and the large-sample approximation indicate that there is strong evidence that tranquilizer does lead to patient improvement, as measured by a reduction in the Hamilton scale factor IV values.

Wilcoxon's signed rank test

An estimate for median:

```
> library(Rfit)
> sort(walsh(z))
 [1] -1.0220 -0.9870 -0.9520 -0.8210 -0.8060 -0.7860 -0.7710 -0.7560 -0.7260
[10] -0.7210 -0.6910 -0.6200 -0.6050 -0.5900 -0.5550 -0.5400 -0.5250 -0.5160
[19] -0.5100 -0.4900 -0.4810 -0.4710 -0.4600 -0.4375 -0.4360 -0.4300 -0.4025
[28] -0.3150 -0.3000 -0.2700 -0.2550 -0.2500 -0.2365 -0.2215 -0.2200 -0.2050
[37] -0.1750 -0.1715 -0.1415 -0.0100  0.0350  0.0685  0.0800  0.1135  0.1470
```

$$M = 45$$

$$\Rightarrow \hat{\theta} = W^{(23)} = -0.46$$

Wilcoxon's signed rank test

Confidence interval for median:

With $n = 9$ and $\alpha = .05$, each configuration under null has equal probability of $\frac{1}{2^9} = 0.001953125$, there should be at most 12.8 configurations to the right of $t_{\alpha/2} = 40$. Thus, $t_{1-\alpha/2} = 45 + 1 - 40 = 6$.

$\theta_{\mathrm{L}} = W^{(6)} = -.786$ and $\theta_{\mathrm{U}} = W^{(40)} = -.010$

so that our 95% confidence interval for θ is

$$(\theta_{\mathrm{L}}, \theta_{\mathrm{U}}) = (-.786, -.010)$$

Wilcoxon's signed rank test

Confirm with built-in function:

```
> wilcox.test(post,pre,paired = T,conf.int = T,conf.level = 0.95)

        Wilcoxon signed rank exact test

data:  post and pre
V = 5, p-value = 0.03906
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -0.786 -0.010
sample estimates:
(pseudo)median
         -0.46
```

Fisher's signed test

Exact test:

$$B = 2$$

$$p - value = P(B \leq 2 | B \sim Bin(9, 1/2)) = 0.08984375$$

Large-sample approximation:

$$B^* = \frac{2 - \left(\frac{9}{2}\right)}{\left(\frac{9}{4}\right)^{1/2}} = -1.666667$$

$$p - value = P(Z < -1.666667) = 0.04779032$$

Both the exact test and the large-sample approximation indicate that there is strong evidence that tranquilizer does lead to patient improvement, as measured by a reduction in the Hamilton scale factor IV values.

same qualitative conclusion

Fisher's signed test

Confirm with built-in function:

```
> SIGN.test(z,alt='less',conf.level = 0.95)

        One-sample Sign-Test

data:  z
s = 2, p-value = 0.08984
alternative hypothesis: true median is less than 0
95 percent confidence interval:
  -Inf 0.041
sample estimates:
median of x
      -0.49
```

Fisher's signed test

An estimate for median:

The ordered Z observations are $Z^{(1)} \leq \cdots \leq Z^{(9)}$:

```
> sort(z)
[1] -1.022 -0.952 -0.620 -0.590 -0.490 -0.430 -0.010  0.080  0.147
```

$$\hat{\theta} = Z^{(5)} = -0.49$$

Fisher's signed test

Confidence interval for median: With $n = 9$ and $\alpha = .05$, the null distribution of B:

```
> dbinom(x=seq(0,9,by=1), size=9, prob=0.5)
 [1] 0.001953125 0.017578125 0.070312500 0.164062500 0.246093750 0.246093750
 [7] 0.164062500 0.070312500 0.017578125 0.001953125
```

$b_{\alpha/2} = 8$, $b_{1-\alpha/2} = 9 + 1 - 8 = 2$

Correction: 2 → 1, 8 → 9

$$\theta_{\mathrm{L}} = Z^{(1)} = -0.952 \text{ and } \theta_{\mathrm{U}} = Z^{(9)} = 0.080$$

so that our 95% confidence interval for θ is

$$(\theta_{\mathrm{L}}, \theta_{\mathrm{U}}) = (-0.952, 0.080)$$

corresponds to Upper Achieved CI.

actual confidence > 95%

Fisher's signed test

Confirm with built-in function:

```
> library(BSDA)
> SIGN.test(z,alt='two.sided',conf.level = 0.95)

        One-sample Sign-Test

data:  z
s = 2, p-value = 0.1797
alternative hypothesis: true median is not equal to 0
95 percent confidence interval:
 -0.9261778  0.0730000
sample estimates:
median of x
      -0.49

Achieved and Interpolated Confidence Intervals:

                  Conf.Level  L.E.pt U.E.pt
Lower Achieved CI     0.8203 -0.6200 -0.010
Interpolated CI       0.9500 -0.9262  0.073
Upper Achieved CI     0.9609 -0.9520  0.080
```