Chap 3 # 1, 3, 6, 18, 22, 25, 27, 29, 46, 63.

See attached R code and output.

1. The data in Table 3.3 are a subset of the data obtained by Kaneto, Kosaka, and Nakao (1967). The experiment investigated the effect of vagal nerve stimulation on insulin secretion. The subjects were mongrel dogs with varying body weights. Table 3.3 gives the amount of immunoreactive insulin in pancreatic venous plasma just before stimulation of the left vagus nerve (X) and the amount measured 5 min after stimulation (Y) for seven dogs. Test the hypothesis of no effect against the alternative that stimulation of the vagus nerve increases the blood level of immunoreactive insulin.

Table 3.3 Blood Levels of Immunoreactive Insulin (μU/ml)

Dog i	before X_i	after Y_i
1	350	480
2	200	130
3	240	250
4	290	310
5	90	280
6	370	1450
7	240	280

Source: A. Kaneto, K. Kosaka, and K. Nakao (1967).

H_0: No effect on blood levels of Immunoreactive insulin from stimulation

H_a: The vagus nerve increases the blood level of immunoreactive insulin

For data with before/after timepoint, use paired Wilcoxon signed-rank test.

From R: $V = 4$, p-val $= 0.960 > 0.05$

$\therefore$ We fail to reject H_0 and conclude that there is no effect from stimulation on increasing the blood levels of immunocreative insulin.

3. Let $T^- = \sum_{i=1}^{n} R_i(1-\psi_i)$, where $\psi_i = 1$ if $Z_i > 0$, and 0 otherwise. Verify directly, or illustrate using the data of Table 3.1, the equation $T^+ + T^- = n(n+1)/2$.

Table 3.1 The Hamilton Depression Scale Factor IV Values

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

Source: D. S. Salsburg (1970).

Given $T^- = \sum_{i=1}^{n} R_i(1-\psi_i)$

for $\psi_i = \begin{cases} 1, & Z_i > 0 \\ 0, & \text{o.w.} \end{cases}$

From Table 3.1, we calculated the following

$Z_i = Y_i - X_i$ rank.

i	Z_i	$\lvert Z_i \rvert$	R_i	ψ_i	$R_i\psi_i$	$1-\psi_i$	$R_i(1-\psi_i)$
1	−0.952	0.952	8	0	0	1	8
2	0.147	0.147	3	1	3	0	0
3	−1.022	1.022	9	0	0	1	9
4	−0.430	0.430	4	0	0	1	4
5	−0.620	0.620	7	0	0	1	7
6	−0.590	0.590	6	0	0	1	6
7	−0.490	0.490	5	0	0	1	5
8	0.080	0.080	2	1	2	0	0
9	−0.010	0.010	1	0	0	1	1
					$T^+ = 5$		$T^- = 40$

$\therefore T^+ + T^- = 5 + 40 = 45$

$\frac{n(n+1)}{2} = \frac{9(10)}{2} = 45$

equal.

Hence, verified.

6. For arbitrary number of observations n, what are the smallest and largest possible values for T^+? Justify your answers.

Given $T^+ = \sum_{i=1}^{n} R_i \psi_i$

$$\psi_i = \begin{cases} 1, & z_i > 0 \\ 0, & z_i < 0. \end{cases}$$

① If $z_i < 0$ for $i = 1, \ldots, n$,

then $\psi_i = 0 \;\; \forall i$

$\therefore$ The smallest possible values for T^+ is 0.

② If $z_i > 0$ for $i = 1, \ldots, n$,

then $\psi_i = 1 \;\; \forall i$.

$\therefore$ The largest possible values for T^+ is $n(1) = n$.

19. Estimate θ for the blood-level data of Table 3.3.

Table 3.3 Blood Levels of Immunoreactive Insulin (μU/ml)

Dog i	X_i	Y_i
1	350	480
2	200	130
3	240	250
4	290	310
5	90	280
6	370	1450
7	240	280

Source: A. Kaneto, K. Kosaka, and K. Nakao (1967).

By R:

```
> owa(x, y)
$owa
 [1]  -70  -30  -25  -15   10   15   20   25   30   30   40   60   70   75   85  100  105  115  130  160  190  505  545  550
560  605  635 1080

$h.l
[1] 80
```

The ordered Walsh avgs shown by \$owa.
The HL estimator $\hat{\theta} = 80$.

22. Verify directly, or illustrate using the data of Table 3.1, that (when there are no ties among the absolute values of the Z's and none of the Z's is zero) T^+ is equal to the number of positive Walsh averages W^+. (See Comment 17.)

Table 3.1 The Hamilton Depression Scale Factor IV Values

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

Source: D. S. Salsburg (1970).

From Q3:

$Z_i = Y_i - X_i$ rank.

i	Z_i	$\|Z_i\|$	R_i	ψ_i	$R_i\psi_i$	$1-\psi_i$	$R_i(1-\psi_i)$
1	−0.952	0.952	8	0	0	1	8
2	0.147	0.147	3	1	3	0	0
3	−1.022	1.022	9	0	0	1	9
4	−0.430	0.430	4	0	0	1	4
5	−0.620	0.620	7	0	0	1	7
6	−0.590	0.590	6	0	0	1	6
7	−0.490	0.490	5	0	0	1	5
8	0.080	0.080	2	1	2	0	0
9	−0.010	0.010	1	0	0	1	1
					$T^+ = 5$		$T^- = 40$

Given Walsh averages as $\frac{Z_i + Z_j}{2}$, $i \le j = 1, \ldots, n$. for each of the $\frac{n(n+1)}{2}$ avgs.

W^+ as the # of positive Walsh avgs,

$T^+ = \sum_{i=1}^{n} R_i \psi_i$ for $\psi_i = \begin{cases} 1, & Z_i > 0 \\ 0, & Z_i < 0. \end{cases}$

By Table 3.1, for a total of $\frac{9(10)}{2} = 45$ avgs, we have the Walsh avgs as: (From R)

```
> owa(x, y)
$owa
 [1] -1.0220 -0.9870 -0.9520 -0.8210 -0.8060 -0.7860 -0.7710 -0.7560 -0.7260 -0.7210 -0.6910 -0.6200 -0.6050 -0.5900
[15] -0.5555 -0.5400 -0.5250 -0.5160 -0.5100 -0.4900 -0.4810 -0.4710 -0.4600 -0.4375 -0.4360 -0.4300 -0.4025 -0.3150
[29] -0.3000 -0.2700 -0.2550 -0.2500 -0.2365 -0.2215 -0.2200 -0.2050 -0.1750 -0.1715 -0.1415 -0.0100  0.0350  0.0685
[43]  0.0800  0.1135  0.1470

$h.l
[1] -0.46
```

$\therefore W^+ = 5 = T^+ = 5.$

25. Explain why the Hodges–Lehmann estimator is less influenced by outlying observations than is the sample mean of the Z's.

The HL estimator is defined by

$$\hat{\theta} = \text{median}\left\{\frac{Z_i + Z_j}{2},\ i \le j = 1, \dots, n\right\}$$

where $\frac{Z_i + Z_j}{2}$ avgs are known as Walsh avgs.

By definition, the HL estimate is estimated from the median of the Walsh avgs that were calculated from $Z_i = Y_i - X_i$ with mutually independent Z_i for $i = 1, \dots, n$.

∴ The HL estimator finds population's median with symmetric distribution and it's robust to unbounded values such as outliers by nature.

But the sample mean $\bar{x} = \frac{\Sigma_i X_i}{n}$, $\bar{y} = \frac{\Sigma_i Y_i}{n}$ is proportional to each single observation and it's strongly affected by outliers.

∴ The HL estimator is less influenced by outliers compared to sample mean.

27. For the blood-level data of Table 3.3, obtain a confidence interval for θ with the exact confidence coefficient .954.

Table 3.3 Blood Levels of Immunoreactive Insulin (μU/ml)

Dog i	X_i	Y_i
1	350	480
2	200	130
3	240	250
4	290	310
5	90	280
6	370	1450
7	240	280

Source: A. Kaneto, K. Kosaka, and K. Nakao (1967).

By R:

```
> wilcox.test(y-x, conf.int = T, conf.level = 0.954)

        Wilcoxon signed rank exact test

data:  y - x
V = 24, p-value = 0.1094
alternative hypothesis: true location is not equal to 0
95.4 percent confidence interval:
 -30 635
sample estimates:
(pseudo)median
            80
```

$\therefore$ The CI for θ with exact confidence coefficient 0.954

$(\theta_L, \theta_U) = (-30, 635)$

29. For the blood-level data of Table 3.3 and $\alpha = .078$, calculate the point estimator of θ defined in Comment 23. Compare with the value of $\hat{\theta}$ obtained in Problem 19.

Table 3.3 Blood Levels of Immunoreactive Insulin (μU/ml)

Dog i	X_i	Y_i
1	350	480
2	200	130
3	240	250
4	290	310
5	90	280
6	370	1450
7	240	280

Source: A. Kaneto, K. Kosaka, and K. Nakao (1967).

By R, with $\alpha = 0.078$, we first calculate the CI:

```
> wilcox.test(y-x, conf.int = T, conf.level = 1 - 0.078)

        Wilcoxon signed rank exact test

data:  y - x
V = 24, p-value = 0.1094
alternative hypothesis: true location is not equal to 0
92.2 percent confidence interval:
 -25 605
sample estimates:
(pseudo)median
            80
```

$\therefore (\theta_L, \theta_U) = (-25, 605)$

By Comment 23, the point estimator of θ is

$$\hat{\theta} = \frac{-25+605}{2} = 290.$$

From problem 19, $\hat{\theta}' = 80$, and $\hat{\theta} > \hat{\theta}'$

46. In an investigation to determine the effect of aspirin on bleeding time and platelet adhesion, Bick, Adams, and Schmalhorst (1976) studied the reactions of normal subjects to aspirin. A subset of their data is presented in Table 3.7, where the X observation for each subject is the bleeding time (in seconds) before ingestion of 600 mg of aspirin and the Y observation is the bleeding time (again in seconds) 2 h after administration of the aspirin.

Perform the appropriate test of the hypothesis that a 600-mg dose of aspirin has no effect on bleeding time versus the alternative that it typically leads to an increase in bleeding time.

Table 3.7 Bleeding Time (in seconds)

Subject i	bef: X_i	after Y_i
1	270	525
2	150	570
3	270	190
4	420	395
5	202	370
6	255	210
7	165	490
8	220	250
9	305	360
10	210	285
11	240	630
12	300	385
13	300	195
14	70	295

Source: R. L. Bick, T. Adams, and W. R. Schmalhorst (1976).

H_0: 600-mg dose of aspirin has no effect on bleeding time

H_a: 600-mg dose of aspirin leads to an increase in bleeding time

By R, use the paired Wilcoxon Signed rank test:

```
> wilcox.test(x, y, paired = T, alternative = "greater")

        Wilcoxon signed rank exact test

data:  x and y
V = 18, p-value = 0.9877
alternative hypothesis: true location shift is greater than 0
```

$V = 18$, $p\text{-val} = 0.9877 > 0.05$.

$\therefore$ We fail to reject H_0 to conclude that 600-mg dose of aspirin has no effect on bleeding time

63. Calculate $\tilde{\theta}$ for the blood-level data of Table 3.3. Compare with the value of $\hat{\theta}$ obtained in Problem 19.

Table 3.3 Blood Levels of Immunoreactive Insulin (μU/ml)

Dog i	X_i	Y_i
1	350	480
2	200	130
3	240	250
4	290	310
5	90	280
6	370	1450
7	240	280

Source: A. Kaneto, K. Kosaka, and K. Nakao (1967).

By (3.58), from R:

```
> z = y-x
> median(z)
[1] 40
```

$\therefore \tilde{\theta} = \text{median}(Z_i,\ 1 \le i \le n) = 40.$

Compared with Prob 19, $\hat{\theta} = 80.$

$\tilde{\theta} < \hat{\theta}$ and $\tilde{\theta} = \frac{1}{2}\hat{\theta}$

ph1855_hw2_ygu5

Yue Gu

2024-01-28

#HW2: Chap3 Q1, 3, 6, 19, 22, 25, 27, 29, 46, 63 ## Page54 Q1

```r
# conduct wilcoxon signed-rank test
# input data
x = c(350, 200, 240, 290, 90, 370, 240)
y = c(480, 130, 250, 310, 280, 1450, 280)

# run the paired wilcoxon signed-rank test
wilcox.test(x, y, paired = T, alternative = "greater")
```

```
##
##  Wilcoxon signed rank exact test
##
## data:  x and y
## V = 4, p-value = 0.9609
## alternative hypothesis: true location shift is greater than 0
```

Page58 Q19

```r
# need to calculate theta-the HL estimator for treatment effects
# first input the data
x = c(350, 200, 240, 290, 90, 370, 240)
y = c(480, 130, 250, 310, 280, 1450, 280)
# calculate theta
# install.packages("NSM3")
library(NSM3)
owa(x, y)
```

```
## $owa
##  [1]  -70  -30  -25  -15   10   15   20   25   30   30   40   60   70   75   85
## [16]  100  105  115  130  160  190  505  545  550  560  605  635 1080
##
## $h.l
## [1] 80
```

Page58 Q22

```r
# input data
x = c(1.83, 0.50, 1.62, 2.48, 1.68, 1.88, 1.55, 3.06, 1.30)
y = c(0.878, 0.647, 0.598, 2.05, 1.06, 1.29, 1.06, 3.14, 1.29)
# calculate Walsh averages
owa(x, y)
```

```
## $owa
##  [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
##
## $h.l
## [1] -0.46
```

Page62 Q27

```r
#input data
x = c(350, 200, 240, 290, 90, 370, 240)
y = c(480, 130, 250, 310, 280, 1450, 280)
# calculate the confidence interval given the exact coefficient
wilcox.test(y-x, conf.int = T, conf.level = 0.954)
```

```
##
##  Wilcoxon signed rank exact test
##
## data:  y - x
## V = 24, p-value = 0.1094
## alternative hypothesis: true location is not equal to 0
## 95.4 percent confidence interval:
##  -30 635
## sample estimates:
## (pseudo)median
##             80
```

Page62 Q29

```r
#input data
x = c(350, 200, 240, 290, 90, 370, 240)
y = c(480, 130, 250, 310, 280, 1450, 280)
# calculate the confidence interval given the exact coefficient
wilcox.test(y-x, conf.int = T, conf.level = 1 - 0.078)
```

```
##
##  Wilcoxon signed rank exact test
##
## data:  y - x
```

```
## V = 24, p-value = 0.1094
## alternative hypothesis: true location is not equal to 0
## 92.2 percent confidence interval:
##  -25 605
## sample estimates:
## (pseudo)median
##             80
```

Page74 Q46

```r
#input data
x = c(270, 150, 270, 420, 202, 255, 165, 220, 305, 210, 240, 300, 300, 70)
y = c(525, 570, 190, 395, 370, 210, 490, 250, 360, 285, 630, 385, 195, 295)
# calculate the confidence interval given the exact coefficient
wilcox.test(x, y, paired = T, alternative = "greater")
```

```
##
##  Wilcoxon signed rank exact test
##
## data:  x and y
## V = 18, p-value = 0.9877
## alternative hypothesis: true location shift is greater than 0
```

Page79 Q63

```r
#input data
x = c(350, 200, 240, 290, 90, 370, 240)
y = c(480, 130, 250, 310, 280, 1450, 280)
#calculate theta-tilt
z = y-x
median(z)
```

```
## [1] 40
```