Module 5 | Assignment: Nonparametric Statistical Methods

Trang Tran

CPS, Northeastern University

ALY6015 | Intermediate Analytics

Professor Steve Morin

May 17, 2023

Section 13-23/ Single-sample Sign Test/ Traditional Method

6. Game Attendance

a. H0: Median = 3000 (claim)

H1: Median ≠ 3000

b. Two-tailed test, α = 0.05; n = 20 => CV = 5 (Table J)

+	+	-	+	+
+	+	-	-	-
-	-	-	+	+
+	-	+	-	-

c. Plus signs = Minus signs = 10 => Test Value = 10

d. Since 10 > 5 => do not reject the null hypothesis.

e. There is not enough evidence to reject the claim that the median number for paid attendance at 20 local football games is 3000.

10. Lottery Ticket Sales

a. H0: Median = 200

H1: Median < 200 (claim)

b. Left-tailed test, α = 0.05; n = 40 => CV = -1.65 (Table E)

c. Test value = z = -1.423 (x = 15)

d. Since -1.423 > -1.65 => Do not reject the null hypothesis.

e. There is not enough evidence to support the claim that the median is below 200 tickets.

Section 13-3/ The Wilcoxon Rank Sum Test/ Traditional Method

4. Lengths of Prison Sentences

a. H0: There is no difference in the sentence received by each gender. (claim)

H1: There is a difference in the length of the sentences of the males and females.

b. Two-tailed test, $\alpha = 0.05$ => CV = ± 1.96 (Table E)

c. Sum of ranks of the group with smaller sample size (males) = 191

$\mu_R = 162$; $\sigma_R = 19.44$

⇨ $z = 1.49$

Length (month)	2	3	4	5	6	7	8	9	11
Gender	F	F	F	F	M	F	M	F	F
Rank	1	2	3	4	5	6	7	8	9

Length (month)	12	12	13	14	15	16	17	19	21
Gender	M	F	M	M	M	F	F	M	F
Rank	10.5	10.5	12	13	14	15	16	17	18

Length (month)	22	23	24	26	26	27	30	32
Gender	M	F	M	F	M	M	F	M
Rank	19	20	21	22.5	22.5	24	25	26

d. Since 1.49 is in between -1.96 and 1.96 => Do not reject the null hypothesis.

e. There is not enough evidence to reject the claim that there is no difference in the sentences received by each gender.

8. Winning Baseball Games/ Traditional Method

a. H0: There is no difference in the number of wins.

H1: There is a difference in the number of wins. (claim)

b. Two-tailed test, $\alpha = 0.05$ => CV = ±1.96 (Table E)

c. Sum of ranks of the group with smaller sample size (NL) = 125

$\mu_R = 132$; $\sigma_R = 16.25$

⇨ z = -0.43

wins	86	88	88	89	89	90	91	91	92
League	AL	AL	NL	NL	AL	NL	NL	AL	NL
Rank	1	2.5	2.5	4.5	4.5	6	7.5	7.5	9

wins	95	95	95	96	96	97	100	100	101
League	AL	AL	NL	NL	NL	AL	AL	NL	NL
Rank	11	11	11	13.5	13.5	15	16.5	16.5	18.5

wins	101	102	104	108	108
League	AL	AL	AL	NL	AL
Rank	18.5	20	21	22.5	22.5

d. Since -0.43 is in between -1.96 and 1.96 => Do not reject the null hypothesis.

e. There is not enough evidence to support the claim that there is a difference in the number of wins.

Section 13-4/ The Wilcoxon Signed-Rank Test/ Traditional Method

5. ws = 13, n = 15, α = 0.01, two-tailed

=> CV = 16 (Table K)

Since 13 < 16 => Reject the null hypothesis.

6. ws = 32, n = 28, α = 0.025, one-tailed

=> CV = 117 (Table K)

Since 32 < 117 => Reject the null hypothesis.

7. ws = 65, n = 20, α = 0.05, one-tailed

=> CV = 60 (Table K)

Since 65 > 60 => Do not reject the null hypothesis.

8. ws = 22, n = 14, α = 0.10, two-tailed

=> CV = 26 (Table K)

Since 22 < 26 => Reject the null hypothesis.

Section 13-5/ Kruskal-Wallis test/ Traditional Method

2. Mathematics Literacy Scores

a. H0: There is no difference in mathematical literacy scores of the three groups.

H1: There is a difference in the mathematical literacy scores of the three groups. (claim)

b. α = 0.05; df = 3 – 1 = 2 => CV = 5.991 (Table G)

c. Sum of the ranks for each group:

Group WH = 24

Group EU = 44

Group EA = 52

⇨ H = 4.16

Score	Group	Rank
381	WH	1
391	EA	2
406	WH	3
411	WH	4
474	WH	5
496	EU	6
510	EU	7
513	EU	8
520	EU	9
523	EA	10
527	WH	11
547	EA	12.5
547	EA	12.5
548	EU	14
549	EA	15

d. Since 4.16 < 5.991 => Do not reject the null hypothesis.

e. There is not enough evidence to support the claim that there is a difference in mathematical literacy scores of the three groups.

Section 13-6/ The Spearman rank correlation coefficient and The Runs test

6. Subway and Commuter Rail Passengers

a. H0: p = 0

H1: $p \neq 0$ (claim)

b. $\alpha = 0.05$, $n = 6$ => CV = 0.886 (Table L)

c. Test value

City	Subway	Rank (X1)	Rail	Rank (X2)	X1-X2	d^2
1	845	6	39	3	3	9
2	494	5	291	6	-1	1
3	425	4	142	5	-1	1
4	313	3	103	4	-1	1
5	108	2	33	1	1	1
6	41	1	38	2	-1	1

- ⇨ Sum(d^2) = 14
- ⇨ $r_s = 0.6$

d. Since 0.6 < 0.886 => Do not reject the null hypothesis.

e. There is no significant relationship between subway and commuter rail trips.

Section 14-3/ Simulation Techniques and Monte Carlo Method

16. Prizes in Caramel Corn Boxes

a. List all possible outcomes: prize 1, prize 2, prize 3, prize 4

b. Determine the probability to each outcome: P1 = P2 = P3 = P4 = ¼ = 0.25

c. Compute the result of the average number of boxes a person needs to buy to get all four prizes.

(Using R) => Result = 8.38

```
> attempts.to.win.per.trial <-
+   random.win.simulation(trials = 40,
+                         target = c(1, 2, 3, 4),
+                         probabilities = NULL)
>
> writeLines(paste("\nThe average number of boxes a person needs to buy to get all four prizes is",
+                  round(sum(attempts.to.win.per.trial) /
+                          length(attempts.to.win.per.trial), 2)))

The average number of boxes a person needs to buy to get all four prizes is 8.38
```

18. Lottery Winner

a. List all possible outcomes: letter B, letter I, letter G

b. Determine the probability of each outcome: P(B) = 0.6; P(I) = 0.3; P(G) = 0.1

c. Compute the result of the average average number of tickets a person must buy to win the prize. (Using R) => Result = 9.87

```
> attempts.to.win.per.trial <-
+   random.win.simulation(trials = 30,
+                         target = c('b', 'i', 'g'),
+                         probabilities = c(0.6, 0.3, 0.1))
>
> writeLines(paste("\nThe average average number of tickets a person must buy to win the prize is",
+                  round(sum(attempts.to.win.per.trial) /
+                          length(attempts.to.win.per.trial), 2)))

The average average number of tickets a person must buy to win the prize is 9.87
```

References

1. Bluman, A. G. (2018). Chapters 13 & 14. In *Elementary statistics Book*. McGraw-Hill.