

6.2 Kruskal-Wallis H-Test

Suppose three or more independent samples are drawn and we would like to test if the location parameters are all equal. Assuming that no information is available about the underlying distributions, we can conduct a nonparametric **Kruskal-Wallis H-test**. Below we describe how this test is carried out.

Suppose measurements for k independent samples of respective sizes n_1, \dots, n_k are available. The hypotheses of interest are $H_0 : \theta_1 = \theta_2 = \dots = \theta_k$ and $H_1 : \theta_i \neq \theta_j$ for some $i \neq j, i, j = 1, \dots, k$. To compute the test statistic, we first pool all observations and assign ranks in increasing order. Tied values are given ranks equal to the average of the ranks that would have been assigned were they not tied. Denote by $R_i, i = 1, \dots, k$, the sum of the ranks in the i -th sample. Let $N = n_1 + \dots + n_k$ be the total number of observations. If no ties exist, then the statistic H is computed according to the formula

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1).$$

Suppose now that the ranking procedure results in m sets of ties. Let $T_i, i = 1, \dots, m$, denote the number of ties in the i -th set. Then the H statistic is computed as

$$H = \left[\frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1) \right] / \left[1 - \frac{\sum_{i=1}^m (T_i^3 - T_i)}{N^3 - N} \right].$$

Note that if no ties occur, then all T_i 's are equal to one, and the second formula is simplified to match the first one.

Once the test statistic is computed, it is compared to the appropriate critical value from Table 6 on page 143 in Dr. Olga's Nonparametric book. If H is larger than or equal to the critical value, the null hypothesis is rejected, and the conclusion is that not all k location parameters are the same. To do the post hoc pairwise comparisons, we use the Wilcoxon rank-sum test.

Example. Returning to our original example with the BMI measurements in three communities, we assign the ranks as summarized in the table.

Community A	29.3	31.4	38.7	33.2	30.3
Rank A	2	5	10	6.5	3
Community B	42.0	39.9	44.5	40.7	38.9
Rank B	14	12	15	13	11
Community C	28.8	36.1	37.5	31.0	33.2
Rank C	1	8	9	4	6.5

We have that $N = 15, n_A = n_B = n_C = 5, R_A = 26.5, R_B = 65, R_C = 28.5, m = 1$, and $T_1 = 2$. The test statistic is

$$H = \frac{\frac{12}{(15)(16)} \left(\frac{26.5^2}{5} + \frac{65^2}{5} + \frac{28.5^2}{5} \right) - 3(16)}{1 - \frac{8-2}{15^3-15}} = 9.411807.$$

From the table, the critical value for $n_A = n_B = n_C = 5$ and $\alpha = 0.01$ is 8. It means that we reject the null at the 1% significant level and conclude that not all location parameters are equal. Below we present SAS and R codes that produce p -values for the Kruskal-Wallis test as well as for the post hoc Wilcoxon rank-sum pairwise tests with two-sided alternatives.

In SAS:

```
/*Kruskal-Wallis Test*/
data cohorts;
  input community $ BMI @@;
cards;
A 29.3 A 31.4 A 38.7 A 33.2 A 30.3
B 42.0 B 39.9 B 44.5 B 40.7 B 38.9
C 28.8 C 36.1 C 37.5 C 31.0 C 33.6
;

proc npar1way data=cohorts Wilcoxon;
  class community;
  var BMI;
  exact;
run;
```

Kruskal-Wallis Test			
Chi-Square	DF	Pr > ChiSq	Exact Pr >= ChiSq
9.4200	2	0.0090	0.0024

Since the p -value is below 0.01, we reject the null and conduct pairwise comparisons using the Wilcoxon rank-sum test.

```
proc npar1way data=cohorts Wilcoxon;
  class community;
  var BMI;
  exact;
  where community in ("A","B");
run;
```

Wilcoxon Two-Sample Test							
Statistic (S)	Z	Pr < Z	Pr > Z	t Approximation		Exact	
				Pr < Z	Pr > Z	Pr <= S	Pr >= S-Mean
15.0000	-2.5067	0.0061	0.0122	0.0167	0.0335	0.0040	0.0079
Z includes a continuity correction of 0.5.							

The p -value is less than 0.01, so the location parameters for the BMI distributions for communities A and B are different.

```
proc npar1way data=cohorts Wilcoxon;
  class community;
  var BMI;
  exact;
  where community in ("A","C");
run;
```

Wilcoxon Two-Sample Test							
Statistic (S)	Z	Pr < Z	Pr > Z	t Approximation		Exact	
				Pr < Z	Pr > Z	Pr <= S	Pr >= S-Mean
26.0000	-0.2089	0.4173	0.8345	0.4196	0.8392	0.4206	0.8413
Z includes a continuity correction of 0.5.							

The p -value is larger than 0.05, so the location parameters for the BMI distributions for communities A and C are the same.

```
proc npar1way data=cohorts Wilcoxon;
  class community;
  var BMI;
  exact;
  where community in ("B","C");
run;
```

Wilcoxon Two-Sample Test							
Statistic (S)	Z	Pr > Z	Pr > Z	t Approximation		Exact	
				Pr > Z	Pr > Z	Pr >= S	Pr >= S-Mean
40.0000	2.5067	0.0061	0.0122	0.0167	0.0335	0.0040	0.0079
Z includes a continuity correction of 0.5.							

The p -value is less than 0.01, so the location parameters for the BMI distributions for communities B and C are different.

In R:

```
BMI.A<- c(29.3, 31.4, 38.7, 33.2, 30.3)
BMI.B<- c(42.0, 39.9, 44.5, 40.7, 38.9)
BMI.C<- c(28.8, 36.1, 37.5, 31.0, 33.2)
```

```
BMI.list<- list(BMI.A, BMI.B, BMI.C)
kruskal.test(BMI.list)
```

Kruskal-Wallis rank sum test

```
data: BMI.list
Kruskal-Wallis chi-squared = 9.4118, df = 2, p-value = 0.009042
```

```
library(exactRankTests)
wilcox.exact(BMI.A, BMI.B, paired=FALSE, alternative="two.sided")
```

Exact Wilcoxon rank sum test

```
data: BMI.A and BMI.B
W = 0, p-value = 0.007937
alternative hypothesis: true mu is not equal to 0
```

```
wilcox.exact(BMI.A, BMI.C, paired=FALSE, alternative="two.sided")
```

Exact Wilcoxon rank sum test

```
data: BMI.A and BMI.C
W = 11.5, p-value = 0.881
alternative hypothesis: true mu is not equal to 0
```

```
wilcox.exact(BMI.B, BMI.C, paired=FALSE, alternative="two.sided")
```

Exact Wilcoxon rank sum test

```
data: BMI.B and BMI.C
W = 25, p-value = 0.007937
alternative hypothesis: true mu is not equal to 0
```

□