


- 
- The **Kruskal-Wallis H Test** is a nonparametric procedure that can be used to compare more than two populations in a **completely randomized design**.
 - All $n = n_1 + n_2 + \dots + n_k$ measurements are jointly ranked (i.e. treat as one large sample).
 - We use the sums of the ranks of the k samples to compare the distributions.

The Kruskal-Wallis H Test

✓ Rank the total measurements in all k samples from 1 to n . Tied observations are assigned average of the ranks they would have gotten if not tied.

✓ Calculate

▪ $T_i =$ rank sum for the i th sample $i = 1, 2, \dots, k$

✓ And the test statistic

$$H = \frac{12}{n(n+1)} \sum \frac{T_i^2}{n_i} - 3(n+1)$$

The Kruskal-Wallis H Test

H_0 : the k distributions are identical versus

H_a : at least one distribution is different

Test statistic: *Kruskal-Wallis H*

When H_0 is true, the test statistic H has an approximate chi-square distribution with $df = k-1$.

Use a right-tailed rejection region or p -value based on the Chi-square distribution.

Example

Four groups of students were randomly assigned to be taught with four different techniques, and their achievement test scores were recorded. Are the distributions of test scores the same, or do they differ in location?

1	2	3	4
65	75	59	94
87	69	78	89
73	83	67	80
79	81	62	88

Teaching Methods

	1	2	3	4
	65 (3)	75 (7)	59(1)	94 (16)
	87 (13)	69 (5)	78 (8)	89 (15)
	73 (6)	83 (12)	67 (4)	80 (10)
	79 (9)	81 (11)	62 (2)	88 (14)
T_i	31	35	15	55

Rank the 16 measurements from 1 to 16, and calculate the four rank sums.

H_0 : the distributions of scores are the same
 H_a : the distributions differ in location

$$\begin{aligned} \text{Statistic: } H &= \frac{12}{n(n+1)} \sum \frac{T_i^2}{n_i} - 3(n+1) \\ &= \frac{12}{16(17)} \left(\frac{31^2 + 35^2 + 15^2 + 55^2}{4} \right) - 3(17) = 8.96 \end{aligned}$$

Teaching Methods

H_0 : the distributions of scores are the same

H_a : the distributions differ in location

$$\begin{aligned}\text{Test statistic: } H &= \frac{12}{n(n+1)} \sum \frac{T_i^2}{n_i} - 3(n+1) \\ &= \frac{12}{16(17)} \left(\frac{31^2 + 35^2 + 15^2 + 55^2}{4} \right) - 3(17) = 8.96\end{aligned}$$

Rejection region: For a right-tailed chi-square test with $\alpha = .05$ and $df = 4 - 1 = 3$, reject H_0 if $H \geq 7.81$.

Reject H_0 . There is sufficient evidence to indicate that there is a difference in test scores for the four teaching techniques.



Key Concepts

I. **Nonparametric Methods**

These methods can be used when the data cannot be measured on a quantitative scale, or when

- The numerical scale of measurement is arbitrarily set by the researcher, or when
- The parametric assumptions such as normality or constant variance are seriously violated.

Key Concepts

Kruskal-Wallis H Test: Completely Randomized Design

1. Jointly rank all the observations in the k samples (treat as one large sample of size n say). Calculate the rank sums, $T_i =$ rank sum of sample i , and the test statistic

$$H = \frac{12}{n(n+1)} \sum \frac{T_i^2}{n_i} - 3(n+1)$$

2. If the null hypothesis of equality of distributions is false, H will be unusually large, resulting in a one-tailed test.
3. For sample sizes of five or greater, the rejection region for H is based on the chi-square distribution with $(k - 1)$ degrees of freedom.

Thank you””” and be
proud of what you
are and what you
do.