

Example 19
Kruskal-Wallis
Comparing more than 2 samples
Nonparametric
Setup

A conservation biologist is recommending the purchasing of a property to be used as a nature reserve. The biologist suspects that the number of endangered orchids differs among the potential sites. The biologist has collected 5 quadrats worth of data from each of the four potential sites.

Sample	Orchids
A	64
A	62
A	63
A	68
A	62
B	61
B	64
B	14
B	58
B	59
C	69
C	63
C	64
C	74
C	56
D	96
D	96
D	90
D	70
D	60

Example 19
Kruskal-Wallis
Comparing more than 2 samples
Solution

1. State your question: Do the number of orchids in the four meadows differ significantly?
 - a. Is it a good scientific question? Definable, measurable, and controllable.
 - b. Identify your population: Population size of orchids
 - c. Identify your dependent variable: Population size
 - d. Identify your independent variable: Meadow
2. State your hypothesis set
 - a. Verbal hypothesis: The meadow influences the number of orchids found therein.
 - b. Statistical hypothesis (H_0 , H_A).
 - H_0 : The number of orchids in the different meadows does not differ significantly.
 - H_A : The number of orchids in the different meadows differs significantly.
 - c. Is your hypothesis set exhaustive? Yes
 - d. Is your hypothesis set exclusive? Yes
3. State your significance level: $\alpha=0.05$
4. Select the appropriate test
 - a. Variable scales
 - i. Dependent variable: Ratio
 - o Converted: Ratio \rightarrow Ordinal
 - ii. Independent variable: Nominal
 - o Converted or transformed? No
 - b. What information is given or available?
 - i. Sample data
 - c. Number of samples: 4
 - d. Are the data paired or unpaired? Unpaired
 - e. What aspect of the variable do you want to compare?
 - i. ~~Central tendency -- Means~~
 - ii. Central tendency -- Medians
 - f. ~~State the test to be used: ANOVA~~
 - i. ~~Are the assumptions of the test met? No~~
 - o ~~Random samples -- Assumed~~
 - o ~~Independent samples -- Assumed~~
 - o ~~Normal Distribution -- Tested using Shapiro-Wilk Test -- Failed~~
 - o ~~Equal Variances -- Tested using Levene's Test -- Passed~~

Shapiro-Wilk normality test

data: OrchidsA\$Orchids
W = 0.8059, p-value = 0.09041

Shapiro-Wilk normality test

data: OrchidsB\$Orchids
W = 0.6586, p-value = 0.003352

Shapiro-Wilk normality test

data: OrchidsC\$Orchids
W = 0.984, p-value = 0.9548

Shapiro-Wilk normality test

data: OrchidsD\$Orchids
W = 0.8421, p-value = 0.1709

Levene's Test for Homogeneity of Variance (center = median)

	Df	F value	Pr(>F)
group	3	0.9224	0.4524

- g. State the test to be used: Kruskal-Wallis
- ii. Are the assumptions of the test met? Yes
- Random samples – Assumed
 - Independent samples – Assumed
 - 3 or more samples – Met
 - Minimum of 5 observations per sample – Met

5. Conduct your sampling
5 quadrats in each meadow

	A	B	C	D
	64	61	69	96
	62	64	63	96
	63	14	64	90
	68	58	74	70
	62	59	56	60
n	5	5	5	5
M	63	59	64	90

6. Graph the data

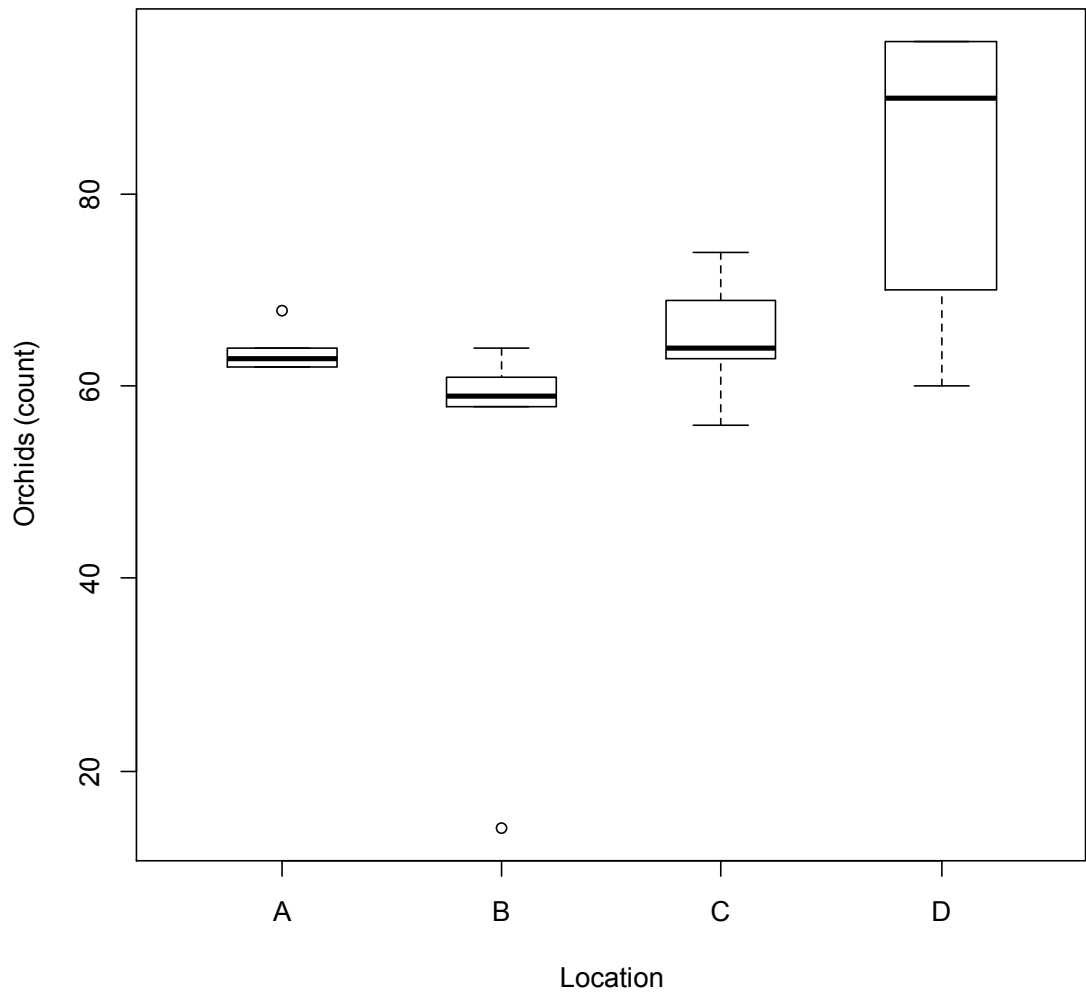


Figure 1. Number of endangered orchids at four potential conservation locations.

7. Summarize the data.

Sample A

$$n_A = 5$$

$$M_A = 63$$

Sample B

$$n_A = 5$$

$$M_A = 59$$

Sample C

$$n_A = 5$$

$$M_A = 64$$

Sample D

$$n_A = 5$$

$$M_A = 90$$

8. Calculate your test statistic.

Kruskal-Wallis rank sum test

data: Orchids\$Orchids by Orchids\$Sample

Kruskal-Wallis chi-squared = 7.8441, df = 3, p-value = 0.04935

9. Retain or reject your null hypothesis based on your test statistic.
The calculated p-value (0.049) is less than the significance level (0.05), so we reject our null hypothesis and retain our alternate hypothesis.
10. Interpret the results in biological terms.
The number of orchids in the different meadows differs significantly ($H=7.844$, $df=3$, $p=0.049$).

Example 19
Kruskal-Wallis
Comparing more than 2 samples
Solution

8. Calculate your test statistic.

Kruskal-Wallis rank sum test

data: Orchids\$Orchids by Orchids\$Sample
 Kruskal-Wallis chi-squared = 7.8441, df = 3, p-value = 0.04935

Multiple comparison test after Kruskal-Wallis

p.value: 0.05

Comparisons

	obs.dif	critical.dif	difference
A-B	4.9	9.871455	FALSE
A-C	1.0	9.871455	FALSE
A-D	5.5	9.871455	FALSE
B-C	5.9	9.871455	FALSE
B-D	10.4	9.871455	TRUE
C-D	4.5	9.871455	FALSE

9. Retain or reject your null hypothesis based on your test statistic.
 The calculated p-value (0.002) is less than the significance level (0.05), so we reject our null hypothesis and retain our alternate hypothesis.

Observation	p
A vs. B	p > 0.05
A vs. C	p > 0.05
A vs. D	p > 0.05
B vs. C	p > 0.05
B vs. D	p < 0.05
C vs. D	p > 0.05

10. Interpret the results in biological terms.
 The number of orchids in the different meadows differs significantly (H=7.844, df=3, p=0.049).