Kruskal-Wallis test (a non-parametric analogue of one-way ANOVA)

The assumptions of the usual one way ANOVA are:

$$X_{ij} = \mu_i + e_{ij}$$

where e_{ij} are independent $N(0, \sigma^2)$, $i = 1, 2, \ldots, a$, $j = 1, 2, \ldots, n_i$

If a normal scores or normal probability plot of the residuals casts doubt on the assumption of normal errors, then one way ANOVA may not give a valid p-value. In such cases, the Kruskal-Wallis test is useful. This is the analogue of the Wilcoxin test for more than two populations.

Kruskal-Wallis test

Assumptions:

- $$X_{ij} = \mu_i + e_{ij}$$

  where e_{ij} are **independent observations from some common distribution.**
- The null hypothesis is $H_0 : \mu_1 = \mu_2 = \ldots = \mu_a$. The alternative is that not all of the means are identical.
- There are $n = \sum_{i=1}^{a} n_i$ total observations.
- Rank the n observations from smallest to largest.
- Let R_{ij} be the rank of the j'th observation in the i'th sample (these ranks will be numbers between 1 and n).
- Let $\bar{R}_{i.}$ be the average rank of the observations in the i'th sample. Let $R_{i.}$ be the sum of the ranks of the observations in the i'th sample.
- The test statistic is

  $$K = \frac{12}{n(n+1)} \sum_{i=1}^{a} n_i \left(\bar{R}_{i.} - \frac{n+1}{2} \right)^2$$

  This is equal to

  $$K = \frac{12}{n(n+1)} \left(\sum_{i=1}^{a} \frac{R_{i.}^2}{n_i} \right) - 3(n+1)$$
- Under the null hypothesis, K has a χ^2 distribution with $a-1$ degrees of freedom, denoted χ^2_{a-1}.
- The p-value is $P(K > K_{obs})$ where K has a χ^2_{a-1} distribution - a chi-squared distribution with $a - 1$ degrees of freedom.
- What is $P(\chi^2_4 > 8)$?
- What is $P(\chi^2_{12} > 24)$?

Example: A group of 32 rats were randomly assigned to each of 4 diets labelled (A,B,C,and D). The response is the liver weight as a percentage of body weight. Two rats escaped and another died, resulting in the following data

A	B	C	D
3.42	3.17	3.34	3.65
3.96	3.63	3.72	3.93
3.87	3.38	3.81	3.77
4.19	3.47	3.66	4.18
3.58	3.39	3.55	4.21
3.76	3.41	3.51	3.88
3.84	3.55		3.96
	3.44		3.91

In minitab, liver weights are in C10, and treatment identifiers are in C11.

```
 C10
   3.42   3.96   3.87   4.19   3.58   3.76   3.84   3.17   3.63   3.38   3.47
   3.39   3.41   3.55   3.44   3.34   3.72   3.81   3.66   3.55   3.51   3.65
   3.93   3.77   4.18   4.21   3.88   3.96   3.91

C11
   1   1   1   1   1   1   1   2   2   2   2   2   2   2   2   3   3   3   3
   3   3   4   4   4   4   4   4   4   4
```

Use the minitab **rank** command to get ranks of the liver weight data. For example, 3.17 is the smallest data value, so it gets rank 1. 3.34 is next smallest value, so gets rank 2, and so on. Note that midranks are used for ties, as in the Wilcoxon test.

```
 MTB > rank c10 c12

C12
    6.0   25.5   21.0   28.0   12.0   17.0   20.0    1.0   13.0    3.0    8.0
    4.0    5.0   10.5    7.0    2.0   16.0   19.0   15.0   10.5    9.0   14.0
   24.0   18.0   27.0   29.0   22.0   25.5   23.0

MTB > let c13=c12**2
```

C12 contains the ranks, and C13 contains the squares of the ranks. Get the sum of the ranks by group.

```
MTB > Describe c12;
SUBC>   By c11;
SUBC>   Sums.

Descriptive Statistics: C12

Variable  C11     Sum
C12       1    129.50
          2     51.50
          3     71.50
          4    182.50

MTB > set c15    C15 contains the sums of the ranks within groups.
DATA> 129.5 51.5 71.5 182.5
DATA> end
```

$$K = \frac{12}{n(n+1)} \sum_{i=1}^{a} n_i \left(\bar{R}_{i.} - \frac{n+1}{2} \right)^2$$

In minitab, this can be done as follows, where C16 contains the sample sizes 7,8,6,8.

```
MTB > let c17=c15/c16
MTB > let k1=(12/(29*30))*sum(c16*(c17-30/2)**2)
MTB > print k1
 K1    16.7945
```

Alternatively,

$$K = \frac{12}{n(n+1)} (\sum_{i=1}^{a} \frac{R_{i.}^2}{n_i}) - 3(n+1)$$

```
MTB > let k2=(12/(29*30))*sum(c15**2/c16)-3*30
MTB > print k2
K2    16.7945
```

- Using either formula, the observed value of the test statistic is $K = 16.8$.
- The p-value is $P(\chi^2_3 > 16.8)$.

```
MTB > cdf 16.8;
SUBC> chisq 3.

Cumulative Distribution Function

Chi-Square with 3 DF

   x  P(X<=x)
16.8     0.999223
```

The p-value is approximately $1 - .999$, or .001.

- To verify the result, we can use the Kruskal-Wallis procedure in minitab.

```
MTB > Kruskal-Wallis c10 c11.

Kruskal-Wallis Test: C10 versus C11

Kruskal-Wallis Test on C10

C11       N  Median  Ave Rank      Z
1         7   3.840      18.5   1.25
2         8   3.425       6.4  -3.34
3         6   3.605      11.9  -1.00
4         8   3.920      22.8   3.05
Overall  29              15.0

H = 16.79  DF = 3  P = 0.001
H = 16.80  DF = 3  P = 0.001  (adjusted for ties)
```

Note the observed value of the test statistic (16.8), and the associated p-value (.001).

- The Krusal-Wallis test indicates that there is very strong evidence against the null hypothesis.