Click on 'Bookmarks' in the Left-Hand menu and then click on the required section.

12.5 Kruskal-Wallis & Friedman Statistics (Web)

12.5.1 Introduction

The Mann-Whitney Test (12.2) looks for differences in median values between *two* samples. Both the Kruskal-Wallis and Friedman Tests look for differences in median values between *more than two* samples.

The Kruskal-Wallis Test is used to analyse the effects of more than two levels of just *one factor* on the experimental result. It is the non-parametric equivalent of the One Way ANOVA (11.1).

The Friedman Test analyses the effect of two factors, and is the non-parametric equivalent of the Two Way ANOVA (11.2).

12.5.2 Kruskal-Wallis Test

Example 12.11

The data in the Table below, gives the efficiency of a chemical process using three different catalysts (A, B and C) on each of four days:

Catalyst	Day 1	Day 2	Day 3	Day 4
A	84.5	82.8	79.1	80.2
B	78.4	79.1	78.0	76.0
C	83.1	79.9	77.8	77.9

Is there evidence that the different catalysts result in different efficiencies?
Assume in this example, that the data may not be normally distributed and that it is necessary to use non-parametric statistics.
(Example 11.2, answers the same problem using the parametric One Way ANOVA)
See following text for calculations:

In Example 12.11, the *factor* involved is the choice of catalyst, which occurs at three *levels*: catalyst A, catalyst B and catalyst C. Each level of the factor is tested with a number of replicate measurements, giving 3 sample sets overall.

The hypotheses for the test are:

Proposed Hypothesis, H_1: Catalysts do have a differential effect on efficiency
Null Hypothesis, H_0: Catalysts have no differential effect on efficiency

The first step is to assign ranks to *all* data values. Set out the data values in order and rank them sequentially, as in Table 12.10. Data items that have equal values are given the *average rank* of those items (see also 12.2.2).

No.	76	77.8	77.9	78	78.4	79.1	79.1	79.9	80.2	82.8	83.1	84.5
Rank	1	2	3	4	5	6.5	6.5	8	9	10	11	12

Table 12.10 Ranking of data for Example 12.11

Each data value is then replaced, in the original data table, by its ranking:

Catalyst	Day 1	Day 2	Day 3	Day 4	**Rank Sum, W_i**
A	12	10	6.5	9	37.5
B	5	6.5	4	1	16.5
C	11	8	2	3	24

Table 12.11 Calculation of rank sums for each sample

The next step is to calculate the sum of the ranks, W_i, for each level, *i*, of the factor - as shown in Table 12.11.

The general test statistic, *H*, for '*r* ' levels of a factor is calculated from the sums, W_i, of each of the *r* samples, i.e. for $i = 1$ to $i = r$:

$$H = \frac{12}{N(N+1)}\left\{\frac{W_1^2}{n_1}+\frac{W_2^2}{n_2}+\frac{W_3^2}{n_3}+.......+\frac{W_r^2}{n_r}\right\}-3(N+1) \qquad [12.3]$$

where n_i is the number of data items in sample, *i*, and *N* is the total number of all data items:

$$N = n_1 + n_2 + n_3 + ... + n_r = \sum_i n_i \qquad [12.4]$$

If there is a significant difference between the medians of different samples, then values for W_i will also be significantly different. It can be shown that such a variation will result in a larger value for the test statistic, *H*.

The Proposed Hypothesis is accepted if $H \geq H_{CRIT}$

For small data samples, the critical value, H_{CRIT}, for the Kruskal-Wallis statistic is dependent on the individual sample sizes, n_i, and the specific values can be found in published tables.

For larger data samples, and as an *approximation* for smaller samples, the critical value, H_{CRIT}, is equal to chi-squared critical value, χ^2_{CRIT}, with degrees of freedom given by:

$$df = r - 1 \qquad [12.5]$$

The value of *H* is calculated for example Example 12.11, using [12.3]:

$$H = \frac{12}{12(12+1)}\left\{\frac{37.5^2}{4}+\frac{16.5^2}{4}+\frac{24^2}{4}\right\} - 3(12+1) = 4.356$$

Degrees of freedom, *df* = 3 -1 = 2, giving an approximate critical value:
$H_{CRIT} \approx \chi^2_{CRIT} = 5.99$

Since $H < H_{CRIT}$ we Do Not Reject the Null Hypothesis. We are unable to detect a significant effect due to the different catalysts.

NB For the data in Example 12.11, the *more accurate* critical value H_{CRIT} = 5.69, which, in this case, leads to the same conclusion.

Q12.7 The fungi species richness was measured on a 10-point scale on four different species of trees over a period of four days, and the results tabulated as below:

	Day 1	Day 2	Day 3	Day 4
Tree 1	6	4	3	3
Tree 2	4	3	3	2
Tree 3	4	2	1	1
Tree 4	2	1	2	1

Perform a Kruskal-Wallis test to investigate whether there is a significant difference in fungi species richness between the trees. Take the measurements on different days as being replicate measurements.

See Q12.4 in the book

12.5.3 Friedman Test

In the Kruskal-Wallis Test, it is assumed that every data item, *within each sample*, are *replicates* recorded under exactly the same conditions, and that the only factor involved is the choice of catalyst for each sample.

However, it is possible that a *second* factor is also influencing the results. Example Example 12.12 considers the same problem as Example 12.11, but with the possibility that the *choice* of day of measurement may also be having some effect. The calculation takes this into account.

Example 12.12
Using the same data as in Example 12.11:

Catalyst	Day 1	Day 2	Day 3	Day 4
A	84.5	82.8	79.1	80.2
B	78.4	79.1	78.0	76.0
C	83.1	79.9	77.8	77.9

Investigate whether the 'day' as well as the 'catalyst' may be a factor in affecting the efficiency of the chemical process.
The data is already **blocked** (15.1.4) so that values appropriate to specific days have been grouped (blocked) within specific columns.
(Example 11.4, answers the same problem using the parametric Two Way ANOVA)
See following text for calculations:

The hypotheses are the same as Example 12.11, but, in Example 12.12, we are now looking for an effect on the efficiency of the process ***by*** the catalyst ***blocked by*** the day of measurement (15.1.4).

In this case, the procedure is to rank the data variables **separately** within each of the (vertical) blocks - see Table 12.12. The next step is to calculate the sum of the ranks, W_i, for each level, *i*, of the factor:

Catalyst	Day 1	Day 2	Day 3	Day 4	**Rank Sum, W_i**
A	3	3	3	3	12
B	1	1	2	1	5
C	2	2	1	2	7

Table 12.12 Calculation of rank sums for each sample for Example 12.12

The general test statistic, *S*, for '*r* ' levels of the factor (rows) and '*c* ' blocks (columns) is calculated from the sums, W_i, of each of the *r* levels, i.e. for $i = 1$ to $i = r$:

$$S = \frac{12}{cr(r+1)}\left\{W_1^2 + W_2^2 + W_3^2 + \ldots\ldots + W_r^2\right\} - 3c(r+1) \qquad [12.6]$$

If there is a significant difference between the medians of different samples, then values for W_i will also be significantly different. It can be shown that this variation will result in a large value for the test statistic, *S*.

The Proposed Hypothesis is accepted if $S \geq S_{CRIT}$

For small data samples, the critical value, S_{CRIT}, for the Friedman statistic is dependent on the number of factor levels, *r*, and the number of blocks, *c*, and the specific values can be found in published tables.

For larger data samples, and as an *approximation* for smaller samples, the critical value, S_{CRIT}, is equal to chi-squared critical value, χ^2_{CRIT}, with degrees of freedom given by:

$$df = r - 1 \qquad [12.7]$$

The value of S is calculated for example Example 12.12, using [12.6]:

$$S = \frac{12}{4 \times 3 \times (3+1)}\left\{12^2 + 5^2 + 7^2\right\} - 3 \times 4 \times (3+1) = 6.5$$

Degrees of freedom, df = 3 -1 = 2, giving an approximate critical value:
$S_{CRIT} \approx \chi^2_{CRIT} = 5.99$
Since $S \geq S_{CRIT}$ we Accept the Proposed Hypothesis.

Blocking the data to take account of the effect of the 'day' has made the test more sensitive to the factor effect of the different catalysts. Compare the use of the Friedman Test with the use of the Two Way ANOVA for the same problem in 11.2.3.

NB For the data in Example 12.12, the *more accurate* critical value S_{CRIT} = 6.50, and we would still Accept the Proposed Hypothesis since $S \geq S_{CRIT}$.

Q12.8 Use the same data as in Q12.4 for fungi species richness:

	Day 1	Day 2	Day 3	Day 4
Tree 1	6	4	3	3
Tree 2	4	3	3	2
Tree 3	4	2	1	1
Tree 4	2	1	2	1

Perform a Friedman test to investigate whether there is a significant difference in fungi species richness between the trees (treatment), while blocking the data by 'day'.

See Q12.5 in the book

12.5.4 Friedman Test - Swapping Factors

In example Example 12.12, there are, effectively, *two* factors - the catalyst and the day. We may choose to re-analyse the data to check principally for the effect of the day, while taking the effect of the catalyst into account - see example Example 12.13:

Example 12.13
Using the same data as in Example 12.12, we can choose to look for an effect on the efficiency of the process ***by*** the day of measurement ***blocked by*** the catalyst.
To do this, we first *transpose* the data so that the *day* becomes the *factor* being tested - the rows of the Table now represent the different factor levels (days):

Catalyst	A	B	C
Day 1	84.5	78.4	83.1
Day 2	82.8	79.1	79.9
Day 3	79.1	78	77.8
Day 4	80.2	76	77.9

See following text for calculations:

The *transposed* problem in Example 12.13 is analysed using the same process as in Example 12.5, but noting the new values:

$c = 3$, $r = 4$ and $df = 4 - 1 = 3$

The calculation of rank sum values is given in Table 12.13

Catalyst	A	B	C	**Rank Sum, W_i**
Day 1	4	3	4	11
Day 2	3	4	3	10
Day 3	1	2	1	4
Day 4	2	1	2	5

Table 12.13 Calculation of rank sums for each sample for Example 12.13

Using [12.6] the value of the test statistic in Example 12.13 is calculated:

$$S = \frac{12}{3 \times 4 \times (4+1)} \left\{ 11^2 + 10^2 + 4^2 + 5^2 \right\} - 3 \times 3 \times (4+1) = 7.4$$

Degrees of freedom: $df = 4 - 1 = 3$, giving an approximate critical value:
$S_{CRIT} \approx \chi^2_{CRIT} = 7.81$
Since $S < S_{CRIT}$ we Do Not Reject the Null Hypothesis.

The Friedman Test has been unable to show that the effect of the 'day' is significant at 0.05. Compare this with the more *powerful* Two Way ANOVA in Example 11.4, which gives a *significant* p-value = 0.022 for the effect of the 'day'.

Q12.9 Use the same data as in Q12.7 and Q12.8 for fungi species richness.

Perform a Friedman test to investigate whether there is a significant difference in fungi species richness between the days (treatment), while blocking the data by 'tree'.

See Q12.6 in book