

1 Multiple comparison of non-normally distributed data with the Kruskal-Wallace test

For data that is not normally distributed, the equivalent test to the ANOVA test (for normally distributed data) is the Kruskal-Wallace test. This tests whether all groups are likely to be from the same population.

```
import numpy as np
from scipy import stats

grp1 = np.array([69, 93, 123, 83, 108, 300])
grp2 = np.array([119, 120, 101, 103, 113, 80])
grp3 = np.array([70, 68, 54, 73, 81, 68])
grp4 = np.array([61, 54, 59, 4, 59, 703])

h, p = stats.kruskal(grp1, grp2, grp3, grp4)

print ('P value of there being a significant difference:')
print (p)
```

OUT:

```
P value of there being a significant difference:
0.013911742382969793
```

If the groups do not belong to the same population, between group analysis needs to be undertaken. One method would be to use repeated Mann-Whitney U-tests, but with the P value needed to be considered significant modified by the Bonferroni correction (divide the required significant level by the number of comparisons being made). This however may be overcautious.