Non-parametric analogue to 1-way anova

There is a non-parametric test for a single factor with more than two levels. This is known as the 'Kruskal-Wallace' test.

- For the model $Y_{ij} = \mu + \tau_i + \epsilon_{ij}$, tests whether any two τ_i differ.
- makes same assumptions as Wilcoxon test
- still based upon ranks.

Kruskal-Wallis

This test is based upon the Kruskal-Wallis test statistic

$$H = \frac{12}{N(N+1)} \Sigma_i^k n_i (R_{.i} - R_{..})^2$$

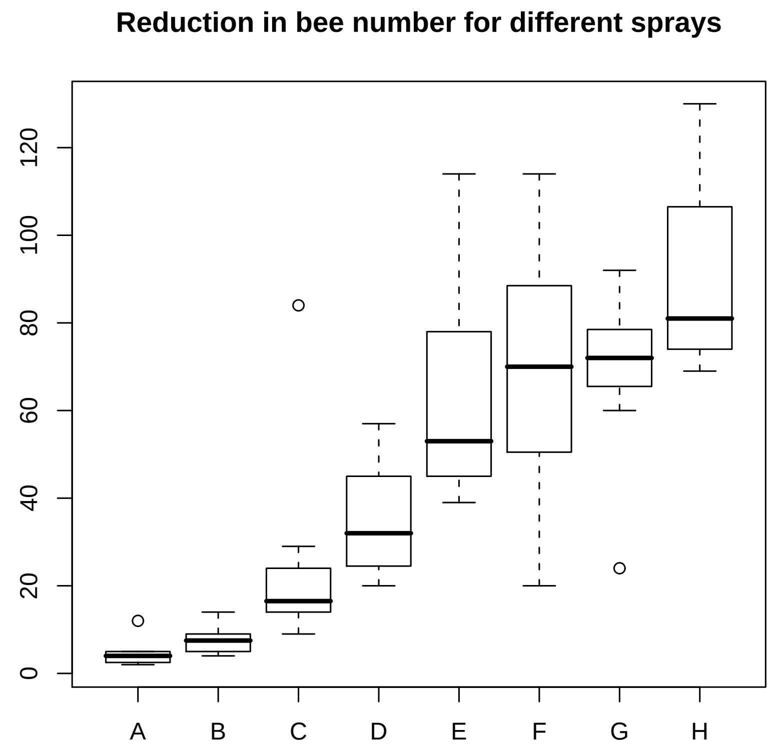
Where $R_{..} = \frac{N+1}{2}$ and $R_{.i}$ is the mean rank in group i.

Kruskal-Wallis, Significance of H

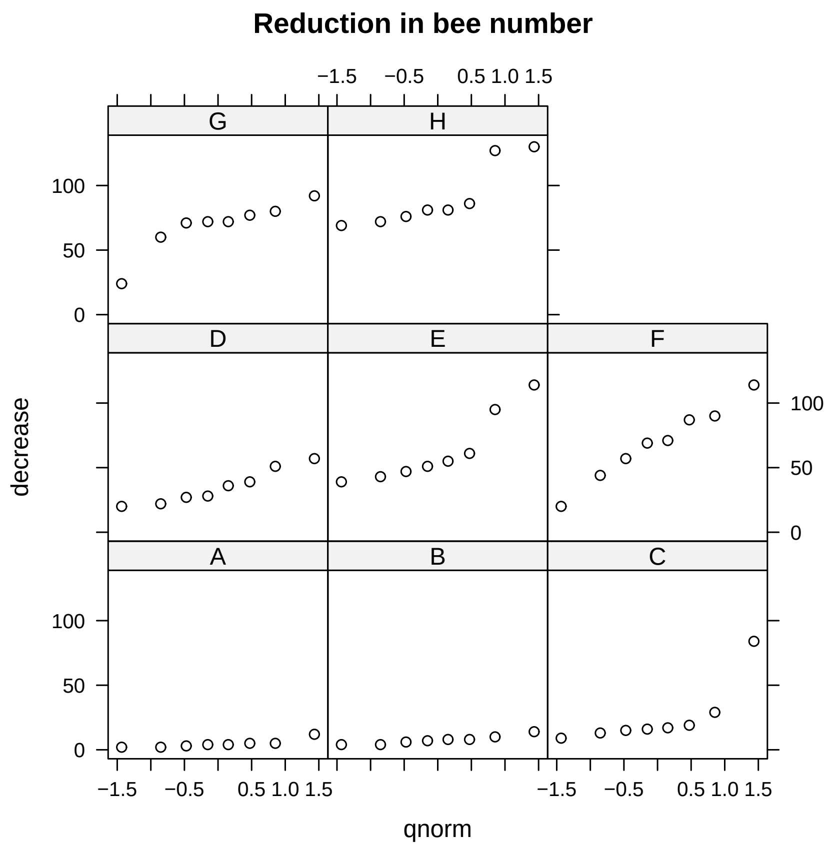
Under H_0, H should have a sampling distribution defined by χ^2 with df $= k - 1$.
If you calculate an H, you can use the R function `pchisq(H,df)` to calculate the probability of observing that value under the null

Kruskal-Wallis

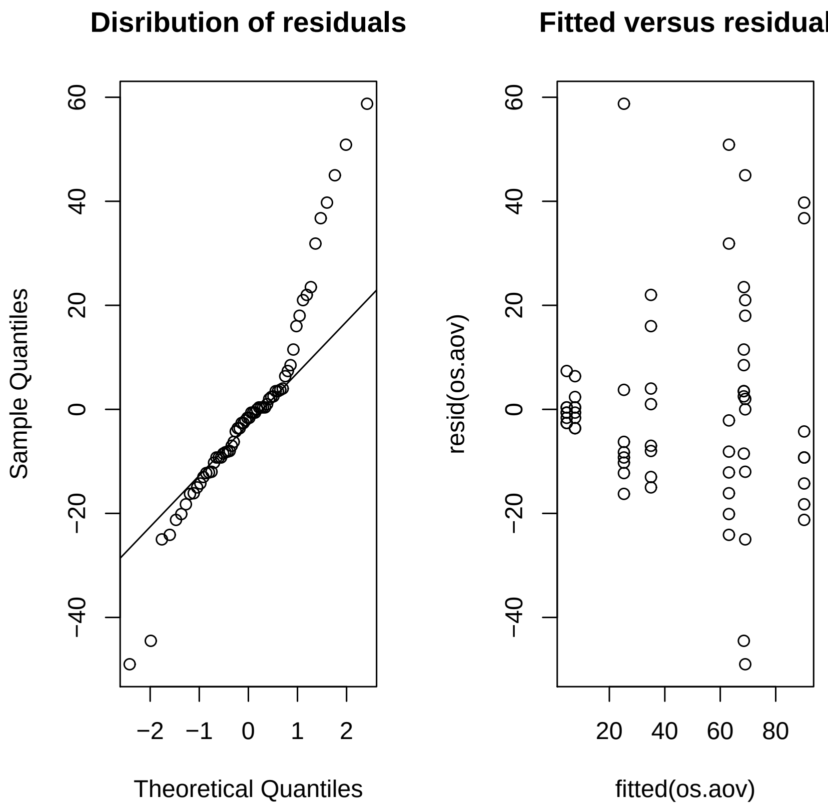
```
> data(OrchardSprays)
> boxplot(decrease ~ treatment, data = OrchardSprays, main = "Reduction in bee number
```



Kruskal-Wallis, cont



Kruskal-Wallis, cont



Kruskal-Wallis, R implementation, cont

```
> library(ctest)
> kruskal.test(decrease ~ treatment, data = OrchardSprays)

        Kruskal-Wallis rank sum test

data:  decrease by treatment
Kruskal-Wallis chi-squared = 48.8735, df = 7, p-value = 2.402e-08
```