One – Way Analysis of Variance (ANOVA)

A common problem encountered in the pharmaceutical industry is when comparisons are required between *m* different groups of data, with responses in each group. The most common experimental design scenario that such data is obtained from is referred to as a completely randomised design; that is one in which the groups are randomly assigned to the experimental units (subjects), or in which the random samples are drawn from each of the *m* populations (groups).

The numbers of observations in each group need not necessarily be equivalent, one can reliably determine if there are differences between the groups through performing a one – way analysis of variance (ANOVA).

The ANOVA procedure evaluates the differences amongst the means of the groups relative to the dispersion in the sampling distributions.

The null hypothesis of interest is that the *m* populations' means are the same, that is:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \ \ldots\ldots = \mu_m$$

$$H_1: \mu_i \text{ are not all equal}$$

A possible method of solving this problem is to run all possible pairwise comparisons of two population means (via *t*-tests), however this proves to be time consuming. Additionally, and more importantly, performing multiple *t*-tests has the consequence of causing the probability of rejecting the null hypothesis to increase. Hence, despite the Type I error being set at $\alpha = 0.05$ for each test, the probability of falsely rejecting at least one of these tests can be considerable larger than this.

ANOVA is a mechanism whereby discriminations between groups are made on the basis of two measurements of variance. One coming from the responses within each group (**within group error**) and the other from differences in group means **(between group differences)**.

The key assumptions of ANOVA are:

- **The subjects have been selected at random from the *m* groups.**
- **The dependent variable/response is normally distributed in each group.**
- **The dependent variable has the same variance in each group.**

If this assumption of **homogeneity of variance** is not met, the statistical test results may not be valid. The term **heteroscedasticity** refers to lack of homogeneity of variances.

The null hypothesis for the test for equality of variances is:

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \ \ldots\ldots = \sigma_m^2$$

$$H_1: \sigma_i^2 \text{ are not all equal}$$

In many experimental contexts, the finding of different variances is as important as the finding of different means. If the variances are truly different, then the populations are different regardless of what ANOVA concludes about differences among the means. **This may be the most important conclusion from the experiment.**

What can we do if the assumption of equal variances is not met?

- Transform the data to equalize the variances, and then re-run the ANOVA.
- Use a modified ANOVA that does not assume equal variances (beyond the scope of this course, this would be **Welch's variance weighted ANOVA**).

Statistically the ANOVA determines whether the group means differ significantly by comparing the two sources of variability to generate an *F*-statistic which is subsequently tested against the corresponding *F*-value.

If the *p*-value is less than 0.05 then we can say that we reject the null hypothesis at the 5% level, however if there are significant differences between the groups it is useful to determine the individual differences between the groups.

3.1 Post – Hoc Comparisons

When the simultaneous comparison of several population means in the ANOVA causes one to reject the null hypothesis of equivalent group means, the logical next step involves determining the specific differences that exist between the groups. This is achieved through the subsequent pairwise comparisons between all of the group means.

In the pairwise techniques that we will be applying, we are only comparing two means at a time. The **Bonferroni test** is probably the most commonly used post hoc test, because it is highly flexible and very simple to compute. Hence, this is the one we will adopt in this course.

Numerous other methods exist for performing post-hoc comparisons in SAS including: Scheffe test, Fisher LSD, and Tukey's HSD for honest significant difference.

3.2 Non – Parametric Alternative

When the assumptions of ANOVA are not met a non-parametric test is required. For a one-way ANOVA one would perform a Kruskal-Wallis test. This is merely an extension of the

Wilcoxon method and the distribution that is used to compare the test statistic against is the X^2 distribution.

The hypothesis that is tested in a Non-Parametric ANOVA is:

H_0: The m distributions are identical

H_1: Not all the distributions are the same

Assumptions:

- the *m* samples have been independently and randomly selected from their respective populations;
- for the chi-square (X^2) approximation there should be >5 observations in each sample;
- tied observations are assigned the average value of the ranks.

As with the post-hoc comparisons, numerous alternatives exist for non-parametric methods for comparing the distribution of groups in an experiment.

Example

Three random samples of laboratory assistants were drawn: one containing 10 academic assistants, the second containing 10 pharmaceutical assistants, and the third containing 10 hospital assistants. Each of the assistants was assessed, to grade their knowledge of various measurement techniques.

The test scores that were obtained were as follows:

Academic	64	61	59	58	52	46	40	38	36	24
Hospital	56	42	31	30	29	28	27	22	18	16
Pharmaceutical	64	62	52	51	44	40	38	34	28	25

This is an example of a case where the Kruskal-Wallis test should be applied because this data are **'scores'**. The data can be found in the file **kruskal_wallis**.

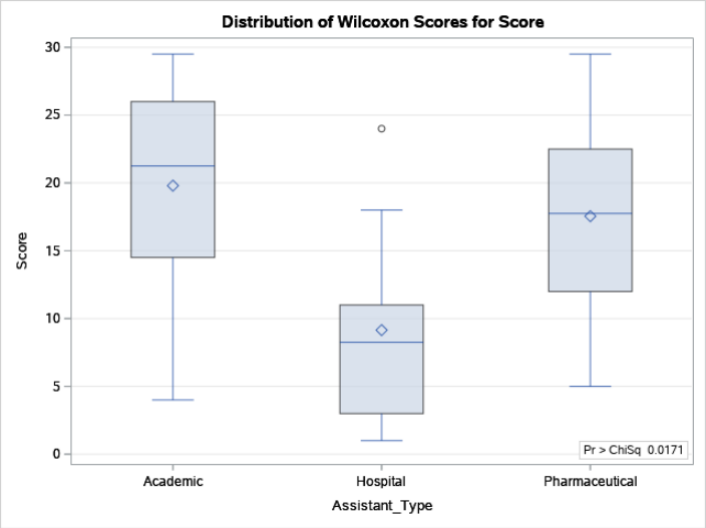
The hypotheses in this case are:

H_o: the distribution of scores for the categories of laboratory assistants are the same;
H_1: the distribution of scores for the categories of laboratory assistants are not all the same.

Using the **Nonparametric One-Way ANOVA (NPAR1WAY)** in SAS. The following table is obtained.

Kruskal-Wallis Test		
Chi-Square	**DF**	**Pr > ChiSq**
8.1400	2	0.0171

It is found that the p-value is estimated as 0.0171, thus we reject the null hypothesis at the 5% level, meaning that at least one of the three groups of assistants differs significantly from the others. To determine which groups differ from which after a significant Kruskal-Wallis test we must examine the post-hoc comparisons along with the visualisations of the scores in each group.



Through selecting the Pairwise multiple comparison analysis option the following table is obtained.

Pairwise Two-Sided Multiple Comparison Analysis			
Dwass, Steel, Critchlow-Fligner Method			
Variable: Score			
Assistant_Type	**Wilcoxon Z**	**DSCF Value**	**Pr > DSCF**
Academic vs. Hospital	2.6458	3.7417	0.0222
Academic vs. Pharmaceutical	0.6057	0.8565	0.8170
Hospital vs. Pharmaceutical	-2.1552	3.0479	0.0791

Based on the p-values, there is a significant difference between Academic and Hospital staff. However, there is no significant difference at the 5% level between Pharmaceutical staff and the 2 other groups. This can also be seen on the box plots above.

We therefore conclude there may be a difference in knowledge of various measurement techniques between hospital assistants and academic assistants, however no significant difference is found with pharmaceutical assistants in any case.