Kruskal-Wallis Non-Parametric ANOVA

Defenition

- The Kruskal–Wallis test is identical to one-way ANOVA with the data replaced by their ranks. It has also been stated that this test is an extension of the two-group Mann-Whitney's U (Wilcoxon rank) test.

Assumptions

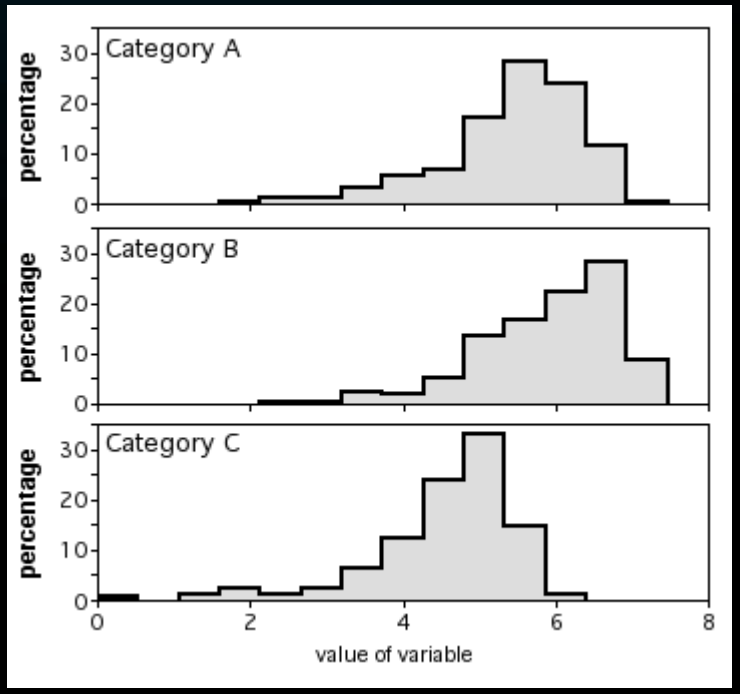
- Assumption #1: Your dependent variable should be measured at the ordinal or continuous level (i.e., interval or ratio).

- Assumption #2: Your independent variable should consist of two or more categorical, independent groups. Typically, a Kruskal-Wallis H test is used when you have three or more categorical, independent groups, but it can be used for just two groups.

- Assumption #3: You should have independence of observations, which means that there is no relationship between the observations in each group or between the groups themselves.

- Assumption #4: In order to know how to interpret the results from a Kruskal-Wallis H test, you have to determine whether the distributions in each group (i.e., the distribution of scores for each group of the independent variable) have the same shape (which also means the same variability).

Assumption 4



Example

- Lymphocyte count determined in four groups in a clinical study is given in Table below:

	Group A	Group B	Group C	Group D
	40.6	31.9	32.7	30.6
	38.0	36.8	31.3	35.9
	41.1	32.4	32.9	29.6
	52.7	34.8	31.9	29.2
	48.8	43.1	28.5	28.5
	41.1	39.0	31.2	30.8
	39.9	33.6	33.1	30.5
	43.1	34.3	34.1	29.4
	32.7	34.0	31.2	30.8
	30.1	33.8	31.7	32.0
Mean	40.8	35.4	31.9	30.7
N	10	10	10	10

Chi-Square Test

Categorical Variable

- We mentioned before that we have two types of variable; categorical and quantitative variables.
- The categorical variable was defined as the frequency of event which might be observed and located under specific title, such as gender, aye colour, Madelaine rate…etc.
- Such variable mostly can’t easily analysed by using the known tests.
- Chi-Square is a test which can be applied in order to highlight the relationships between the these categorical variables.

Chi-Square test

- which is a statistical test used to compare expected data with what we collected.
- i.e
- What a chi-square will tell us is if there is a large difference between collected numbers and expected numbers. If the difference is large, it tells us that there may be something causing a significant change. A significantly large difference will allow us to reject the null hypothesis, which is defined as the prediction that there is no interaction between variables. If the scores are too close, then we have to conclude that they are basically the same.

Chi-Square equation

- Chi-Square denoted by X^2
- The equation which be applied to calculate the Ch-Square is:

$$X^2 = \sum(o-e)^2 / e$$

O= observed value

e= expected value

- Suppose that the ratio of male to female students in the Science Faculty is exactly 1:1, but in the Pharmacology Honours class over the past ten years there have been 80 females and 40 males. Is this a significant departure from expectation?

	Female	Male	Total
Observed numbers (O)	80	40	120
Expected numbers (E)	60	60	120
O - E	20	-20	0
$(O-E)^2$	400	400	
$(O-E)^2 / E$	6.67	6.67	$13.34 = X^2$

Homework

Now repeat this analysis, but knowing that 33.5% of all students in the Science Faculty are males.

If We have two or more categorical variable

- In survey was done by a social researcher to find out the driving aggressive behaviour of both male and female that by giving answer (yes or no). He obtained the following results.
-

gender	aggressive-driving	frequency
male	yes-aggressive	25
male	No-not aggressive	75
female	yes-aggressive	10
female	No-not aggressive	90