UNIT 3 KRUSKAL WALLIS ANALYSIS OF VARIANCE

Structure

3.0 INTRODUCTION

So far in Unit 2 we have studied appropriate statistical tests when we wish to compare two groups (t test if data is from a normal population, Mann-Whitney U test or Wilcoxon test if there are no assumptions about the distribution of the data), but what if there are more than two groups that require comparison? One may think that we may apply the same tests in that condition too. Like for example, if there are three groups say A, B, and C, one may see the difference between A&B, B&C and A&C. This may not look so cumbersome. Now, think if we need to compare 5 groups, A, B, C, D, E, the number for comparison tests we need to do would be 10 (A&B, A&C, A&D, A&E, B&C, B&D, B&E, C&D, C&E, D&E). And what if we need to compare 6 groups? Number of two sample test in these cases become too cumbersome and may not be feasible at all. This may further lead to unnecessary calculations and also give rise to type I error. The answer in these cases when we have more than two groups (>2 groups) to be compared is to conduct Analysis of Variance.

3.0 OBJECTIVES

After reading this unit, you will be able to:

- Define ANOVA tests;
- Describe the procedure for ANOVA calculations;
- Explain Kruskal Wallis ANOVA;
- Enumerate the conditions when this test can be applied; and
- Analyse Kruskal Wallis Anova with one way ANOVA of parametric test.

3.2 ANALYSIS OF VARIANCE

The term analysis of variance (for which the acronym ANOVA is often employed) describes a group of inferential statistical procedures developed by the British statistician Sir Ronald Fisher. Analysis of variance is all about examining the amount of variability in a *y* (response) variable and trying to understand where that variability is coming from. One way that you can use ANOVA is to compare several populations regarding some quantitative variable, *y*. The populations you want to compare constitute different groups (denoted by an *x* variable), such as political affiliations, age groups, or different brands of a product. ANOVA is also particularly suitable for situations involving an experiment where you apply certain treatments *(x)* to subjects, and you measure a response *(y)*.

Null hypothesis H_o%: Population means are equal. There will be no difference in the population means.

ì1 = ì2 = ì3 = ì4

Alternative hypothesis: *h*1

Population means are not equal. There will be difference in the means of the different populations.

The logic used in ANOVA to compare means of multiple groups is similar to that used with the t-test to compare means of two independent groups. When one way ANOVA is applied to the special case of two groups, this one way ANOVA gives identical results as the t-test.

Not surprisingly, the assumptions needed for the t-test are also needed for ANOVA. We need to assume:

1) random, independent sampling from the k populations;

2) normal population distributions;

3) equal variances within the k populations.

Assumption 1 is crucial for any inferential statistic. As with the t-test, Assumptions 2 and 3 can be relaxed when large samples are used, and Assumption 3 can be relaxed when the sample sizes are roughly the same for each group even for small samples. (If there are extreme outliers or errors in the data, we need to deal with them first.)

Self Assessment Questions

1) Fill in the blanks

   i) We would use _____, if we are testing a hypothesis of ì1 = ì2 and _______Test when ì1 = ì2= ì3 = ì4 if the populations under consideration are normally distributed.

   ii) ANOVA was developed by British Statistician __________________.

   iii) ANOVA is used when k __________.

   iv) ANOVA compares multiple means but the logic behind ANOVA is similar to _______ test that compares two independent means.

2) What are the assumptions of ANOVA?

   ....

....

....

....

3) Why are multiple *t*-tests not preferred when we have to compare more than 2 means?

....

....

....

....

3.3 INTRODUCTION TO KRUSKAL-WALLIS ANOVA TEST

When there are more than two groups or k number of groups to be compared, ANOVA is utilised but, again since ANOVA is a parametric statistics and requires assumption of normality as a key assumption, we need to also be aware of its non-parametric counterpart. The Kruskal-Wallis test compares the medians of several (more than two) populations to see whether they are all the same or not. The Kruskal Wallis test is a non-parametric analogue to ANOVA. It can be viewed as ANOVA based on rank transformed data.

That is, the initial data are transformed to their associated ranks before being subjected to ANOVA. In other words, it's like ANOVA, except that it is computed with medians and not means. It can also be viewed as a test of medians.

The null and alternative hypotheses may be stated as:

*H*0: the population medians are equal

*H*1: the population medians differ

3.4 RELEVANT BACKGROUND INFORMATION ON KRUSKAL WALLIS ANOVA TEST

The Kruskal-Wallis one way analysis of variance by ranks (Kruskal, 1952) and (Kruskal and Wallis, 1952) is employed with ordinal (rank order) data in a hypothesis testing situation involving a design with two or more independent samples. The test is an extension of the Mann-Whitney *U* test (Test 12) to a design involving more than two independent samples and, when k = 2, the Kruskal-Wallis one way analysis of variance by ranks will yield a result that is equivalent to that obtained with the Mann-Whitney *U* test.

If the result of the Kruskal-Wallis one-way analysis of variance by ranks is significant, it indicates there is a significant difference between at least two of the sample medians in the set of k medians. As a result of the latter, the researcher can conclude there is a high likelihood that at least two of the samples represent populations with different median values.

In employing the Kruskal-Wallis one-way analysis of variance by ranks one of the following is true with regard to the rank order data that are evaluated:

a) The data are in a rank-order format, since it is the only format in which scores are available; or

b) The data have been transformed into a rank-order format from an interval/ratio format, since the researcher has reason to believe that one or more of the assumptions of the single-factor between-subjects analysis of variance (which is the parametric analog of the Kruskal-Wallis test) are saliently violated.

It should be noted that when a researcher decides to transform a set of interval ratio data into ranks, information is sacrificed. This latter fact accounts for why there is reluctance among some researchers to employ non-parametric tests such as the Kruskal Wallis oneway analysis of variance by ranks, even if there is reason to believe that one or more of the assumptions of the single factor between subjects analysis of variance have been violated.

Various sources {e.g., Conover (1980, 1999), Daniel (1990), and Marascuilo and McSweeney (1977)} note that the Kruskal Wallis one way analysis of variance by ranks is based on the following assumptions:

a) Each sample has been randomly selected from the population it represents;

b) The k samples are independent of one another;

c) The dependent variable (which is subsequently ranked) is a continuous random variable. In truth, this assumption, which is common to many non-parametric tests, is often not adhered to, in that such tests are often employed with a dependent variable which represents a discrete random variable; and

d) The underlying distributions from which the samples are derived are identical in shape.

The shapes of the underlying population distributions, however, do not have to be normal.

Maxwell and Delaney (1990) point out that the assumption of identically shaped distributions implies equal dispersion of data within each distribution. Because of this, they note that, like the single factor between subjects analysis of variance, the Kruskal Wallis one way analysis of variance by ranks assumes homogeneity of variance with respect to the underlying population distribution. Because the latter assumption is not generally acknowledged for the Kruskal Wallis one way analysis of variance by ranks, it is not uncommon for sources to state that violation of the homogeneity of variance assumption justifies use of the Kruskal Wallis one way analysis of variance by ranks in lieu of the single factor between subjects analysis of variance.

It should be pointed out, however, that there is some empirical research which suggests that the sampling distribution for the Kruskal Wallis test statistic is not as affected by violation of the homogeneity of variance assumption as is the F distribution (which is the sampling distribution for the single-factor between-subjects analysis of variance).

One reason cited by various sources for employing the Kruskal Wallis one way analysis of variance by ranks is that by virtue of ranking interval/ratio data a researcher can reduce or eliminate the impact of outliers. As noted earlier in t test for two independent samples, since outliers can dramatically influence variability, they can be responsible for heterogeneity of variance between two or more samples. In addition, outliers can have a dramatic impact on the value of a sample mean.

Zimmerman and Zumbo (1993) note that the result obtained with the Kruskal-Wallis one-way analysis of variance by ranks is equivalent (in terms ofthe derived probability value) to that which will be obtained if the rank-orders employed for the Kruskal-Wallis test are evaluated with a single-factor between-subjects analysis of variance.

Self Assessment Questions

4) Fill in the balnks:

   a) ANOVA is a parametric statistics its equivalent non-parametric statistics is ________________

   b) Kruskal Wallis ANOVA was developed by ____________ and ____________ in 1952.

   c) ANOVA compares means of more than two groups whereas ________ of more than two groups is compared by Kruskal Wallis ANOVA.

   d) One of the assumptions in Kruskal Wallis AONA is that the dependent variable (which is subsequently ranked) is a ___________________

   e) Kruskal Wallis ANOVA can be viewed as ANOVA based on _________ transformed data.

5) State the null and alternative hypothesis for Kruskal Wallis ANOVA.

....

....

....

....

6) Enumerate the assumptions of Kruskal Wallis ANOVA

....

....

....

....

3.5 STEP BY STEP PROCEDURE FOR KRUSKAL WALLIS ANOVA

1) Rank all the numbers in the entire data set from smallest to largest (using all samples combined); in the case of ties, use the average of the ranks that the values would have normally been given.

2) Total the ranks for each of the samples; call those totals $T1$, $T2$, . . ., Tk, where k is the number of groups or populations.

3) Calculate the Kruskal-Wallis test statistic,

   H = [12 / N (N+1)] [$\Sigma((\Sigma R)^2$ / n)] – 3(N + 1)

   N = the total number of cases

   n = the number of cases in a given group

   $(\Sigma R)^2$ = the sum of the ranks squared for a given group of subjects

4) Find the *p*-value.

5) Make your conclusion about whether you can reject Ho by examining the *p*-value.

Example of a Small Sample:

In a Study, 12 participants were divided into three groups of 4 each, they were subjected to three different conditions, A (Low Noise), B(Avearge Noise), and C(Loud Noise). They were given a test and the errors committed by them on the test were noted and are given in the table below.

Participant No.	Condition A (Low Noise)	Participant No.	Condition B (Average Noise)	Participant No.	Condition C (Loud Noise)
1	3	5	2	9	10
2	5	6	7	10	8
3	6	7	9	11	7
4	3	8	8	12	11

The researcher wishes to know whether these three conditions differ amongst themselves. and there are no assumptions of the probability. To apply Kruskal Wallis test, following steps would be taken:

Step 1: Rank all the numbers in the entire data set from smallest to largest (using all samples combined); in the case of ties, use the average of the ranks that the values would have normally been given.

Condition A	Ranks T 1	Condition B	Ranks T2	Condition C	Ranks T3
3	2.5	2	1	10	11
5	4	7	6.5	8	8.5
6	5	9	10	7	6.5
3	2.5	8	8.5	11	12
	ΣT1 = 14		ΣT2 = 26		ΣT3 = 38

Step 2: Total the ranks for each of the samples; call those totals $T1$, $T2$, . . ., Tk, where k is the number of populations.

T1 =14

T2 =26

T3=38

Step3: Caculate H

$H = [\ 12 / N\ (N+1)\ ]\ [\ \Sigma((\Sigma R)^2 / n)\ ] - 3(N + 1)$

N = 12

n = 4

$(\Sigma R)^2 = (14+ 26+ 38)^2 = 6084$

H= [12/ 12 (12 + 1)] [$(14^2/4) + (26^2/4) + (38^2/4)$] – 3 (12+ 1)

H= [12/156] [49 + 169 + 361] – 39

H= (0.076 x 579) – 39

H= 44.525 – 39

H= 5.537

Step 4: Find the *p*-value.

Since the groups are three and number of items in each group are 4, therefore looking in table H (k=3, sample size of 4,4,4) it can be seen that the critical value is 5.692 (α = 0.05).

Step 5: Make your conclusion about whether you can reject Ho by examining the *p*-value.

Since the critical value is more than the actual value we *accept the null hypothesis* that all the three conditions A (Low Noise), B(Avearge Noise), and C(Loud Noise), do not differ from each other, therefore, in the said experiment there was no differences in the groups performance based on the noise level.

3.6 CONSIDERATIONS FOR LARGE SAMPLE

When the number of sample increases, the table H is unable to give us with the critical values, like for example it gives critical values up to 8 samples when k=3, 4 when k=4, and 3 samples when k=5, therefore as the sample increases table H is not of use for the critical value. In such a case we resort to chi square table for getting our information on the critical value taking degrees of freedom (k – 1).

Exact tables of the Kruskal-Wallis distribution: Although an exact probability value can be computed for obtaining a configuration of ranks which is equivalent to or more extreme than the configuration observed in the data evaluated with the Kruskal-Wallis one-way analysis of variance by ranks, the chi-square distribution is generally employed to estimate the latter probability. As the values of *k* and *N* increase, the chi-square distribution provides a more accurate estimate of the exact Kruskal-Wallis distribution. Although most sources employ the chi-square approximation regardless of the values of *k* and *N*, some sources recommend that exact tables be employed under certain conditions. Beyer (1968), Daniel, and Siegel and Castellan (1988) provide exact Kruskal-Wallis probabilities for whenever *k*= 3 and the number of subjects in any of the samples is five or less. Use of the chi-square distribution for small sample sizes will generally result in a slight decrease in the power of the test (i.e., there is a higher likelihood of retaining a false null hypothesis). Thus, for small sample sizes the tabled critical chi-square value should, in actuality, are a little lower than the value of Table H.

Worked Example for a large sample

A state court administrator asked the 24 court coordinators in the state's three largest counties to rate their relative need for training in case flow management on a Likert scale (1 to 7).

1 = no training need

7 = critical training need

Training Need of Court Coordinators

County A	County B	County C
3	7	4
1	6	2
3	5	5
1	7	1
5	3	6
4	1	7
4	6	
2	4	
	4	
	5	

Step 1: Rank order the total groups' Likert scores from lowest to highest.

If tied scores are encountered, sum the tied positions and divide by the number of tied scores. Assign this rank to each of the tied scores.

Scores & Ranks Across the Three Counties

Ratings	Ranks	Ratings	Ranks
1	2.5	4	12
1	2.5	4	12
1	2.5	5	16.5
1	2.5	5	16.5
2	5.5	5	16.5
2	5.5	5	16.5
3	8	6	20
3	8	6	20
3	8	6	20
4	12	7	23
4	12	7	23
4	12	7	23

Calculating the ranks of tied scores

Example: Three court administrators rated their need for training as a 3. These three scores occupy the rank positions 7, 8, & 9.

$$(7 + 8 + 9) / 3 = 8$$

Step 2 Sum the ranks for each group and square the sums

County A		County B		County C	
Rating	Rank	Rating	Rank	Rating	Rank
3	8	7	23	4	12
1	2.5	6	20	2	5.5
3	8	5	16.5	5	16.5
1	2.5	7	23	1	2.5
5	16.5	3	8	6	20
4	12	1	2.5	7	23
4	12	6	20		
2	5.5	4	12		
		4	12		
		5	16.5		
Σ R $(\Sigma R)^2$	67.0 4489		153.5 23562.25		79.5 6320.25

Step 3 Calculate H

$H = [\ 12 / N\ (N+1)\ ]\ [\ \Sigma((\Sigma R)^2 / n)\ ] - 3(N + 1)$

$H = [\ 12 / 24\ (24+1)\ ]\ \ [4489 / 8 + 23562.25 / 10 + 6320.25 / 6] - 3\ (24 + 1)$

$H = (0.02)\ (3970.725) - (75)$

$H = 4.42$

$df = (k - 1) = (3 - 1) = 2$

Interpretation

The critical chi-square table value of H for $\alpha = 0.05$, and df = 2, is 5.991

Since 4.42 < 5.991, the null hypothesis is accepted. There is *no difference* in the training needs of the court coordinators in the three counties

Self Assessment Questions

1) Rearrange the following steps of Kruskal-Wallis test in appropriate order:

   i) Calculate H

   ii) Make your conclusion about whether you can reject Ho by examining the *p*-value.

   iii) Rank all the numbers in the entire data set from smallest to largest

   iv) Find the *p*-value.

   v) Total the ranks for each of the samples; call those totals *T*1, *T*2, . . ., *Tk*, where *k* is the number of populations.

2) Fill in the Blanks

   i) As the values of *k* and *N* increase, the __________ distribution provides a more accurate estimate of the exact Kruskal-Wallis distribution.

   ii) Use of the chi-square distribution for small sample sizes will generally result in a slight _____________ in the power of the test.

   iii) When the critical value of H is more than the actual obtained value of H, we _____________ the null hypothesis.

   iv) When the critical value of H is less than the actual obtained value of H, we _____________ the null hypothesis.

3.7 COMPARISON OF ANOVA AND KRUSKAL WALLIS ANOVA TEST

The Kruskal-Wallis (KW) ANOVA is the non-parametric equivalent of a one-way ANOVA. As it does not assume normality, the KW ANOVA tests the null hypothesis of no difference between three or more group medians, against the alternative hypothesis that a significant difference exists between the medians. The KW ANOVA is basically an extension of the Wilcoxon-Mann-Whitney (WMW) 2 sample test, and so has the same assumptions: 1) the groups have the same spreads; and 2) the data distributions have the same shape.

ANOVA compares means of different population to indicate the similarity between the populations KW ANOVA compares medians of these populations. ANOVA compares the data itself, KW ANOVA, converts data into ranks and then does its computation, in this respect Kruskal Wallis ANOVA is knows as Kruskal Wallis Analysis of Variance by Ranks.

Let's look at the Example to see how their calculations differ:

Three groups 1, 2, and 3, performed a task, we want to see whether they differ or not.

Group 1	Group 2	Group 3		Group 1 (Ranks)	Group 2 (Ranks)	Group 3 (Ranks)
3	6	9		1	3.5	10
5	7	10		2	5.5	13
6	8	11		3.5	7.5	15.5
7	9	12		5.5	10	17
8	10	15		7.5	13	18
9	10			10	13	
	11				15.5	
38	61	57	Total	29.5	68	73.5

ANOVA

$n_1 = 6$ $\quad$ $n_2 = 7$ $\quad$ $n_3 = 5$

$\Sigma X_1 = 38$ $\quad$ $X_2 = 61$ $\quad$ $\Sigma X_3 = 57$

$\Sigma(X_1)^2 = 264$ $\quad$ $\Sigma (X_2)^2 = 551$ $\quad$ $\Sigma (X_3)^2 = 671$

$SS_{total} = (264 + 551 + 671) - [(38 + 61 + 57)^2 / 18] = 134$

$$SS_{Between\ Groups} = (38^2/6) + (61^2/7) + (57^2/5) – [(38 + 61+ 57)^2 / 18]$$

$$SS_{Within\ Gropus} = [\ 264 - (38^2/6)] + [\ 551- (61^2/7)] + [\ 671- (57^2/5)] = 63.962$$

Source of Variation	SS	df	MS	F ratio	F critical Value	Test Decision
Between Groups	70.038	2	35.019	8.213	3.68	Reject H_0
Within Groups	63.962	15	4.264			
Total	134.000	17				

Kruskal Wallis H test:

$H = [12 / 18(18+1)]\ [\ (29.5^2/6) + (68^2/7) + (73.5^2/5)] – [3\ (18+1)]$

$H= 66.177 – 57 = 9.177$

Chi Square for Degrees of freedom 2 (3 – 1) is 5.99, Therefore reject the H_o

In both the case, ANOVA or Kruskal Wallis ANOVA, we will reject the Null Hypothesis, and state that the three groups differ.

F ratio as a function of H:

The fisher's F or F ratio or ANOVA one way variance is equivalent to H test or Kruskal Wallis test or Kruskal Wallis ANOVA, or ANOVA by rank order. This can also be seen in book Iman and Conover (1981)

Where the rank transform statistics states:

$F= [\{(k-1)/(N-k)\}\ \{((N-1)/H)-1\}]^{-1}$

If We see from the above mentioned example

F was 8.213 and H was 9.177

$F= [\{(3-1)/(18-3)\}\ \{((18-1)/9.177)-1\}]^{-1}$

$F= [(2/15)\ \{(17/9.177)-1\}]^{-1}$

$F= [0.133 \text{ x } (1.852-1)]^{-1} = 0.1214^{-1}$

$F= 8.231$

3.8 LET US SUM UP

Kruskal Wallis One way Analysis of Variance (KW ANOVA) is a Non-parametric Analogue of ANOVA one way variance for independent sample. Kruskal Wallis ANOVA is used when there are more than 2 groups (k > 2). The assumption of KW ANOVA are that: a) Each sample has been randomly selected from the population it represents; b) The k samples are independent of one another; c) The dependent variable (which is subsequently ranked) is a continuous random variable. In truth, this assumption, which is common to many non-parametric tests, is often not adhered to, in that such tests are often employed with a dependent variable which represents a discrete random variable; and d) the underlying distributions from which the samples are derived are identical in shape.

The ANOVA or F test and KW ANOVA Or H test are equivalent to each other and can be appropriately used depending upon the type of population in question.

3.9 UNIT END QUESTIONS

1) Under what circumstances does the chi-square distribution provide an appropriate characterisation of the sampling distribution of the Kruskal–Wallis *H* statistic?

2) Data were collected from three populations—*A*, *B*, and *C*,—by means of a completely randomized design.

   The following describes the sample data:

   nA = nB = nC = 15

   RA = 235 RB = 439 RC = 361

   a) Specify the null and alternative hypotheses that should be used in conducting a test of hypothesis to determine whether the probability distributions of populations *A*, *B*, and *C* differ in location.

   b) Conduct the test of part a.

3) A firm wishes to compare four programs for training workers to perform a certain manual task. Twenty new employees are assigned to the training programs, with 5 in each program. At the end of the training period, a test is conducted to see how quickly trainees can perform the task. The number of times the task is performed per minute is recorded for each trainee, with the following results:

Observation	*Program 1*	*Program 2*	*Program 3*	*Program 4*
1	9	10	12	9
2	12	6	14	8
3	14	9	11	11
4	11	9	13	7
5	13	10	11	8

Calculate H, and report your results appropriately

4) An economist wants to test whether mean housing prices are the same regardless of which of 3 air-pollution levels typically prevails. A random sample of house purchases in 3 areas yields the price data below.

Mean Housing Prices (Thousands of Dollars):

MEAN HOUSING PRICES (THOUSANDS OF DOLLARS): *Pollution Level*			
Observation	Low	Mod	High
1	120	61	40
2	68	59	55
3	40	110	73
4	95	75	45
5	83	80	64

Calculate H and report your results with the p-value of 0.05

5) Show that H is equivalent to the F test statistics in one way analysis of variance problem if applied to the ranks of the observation rather than the actual numbers. (**Hint**: Express the F ratio as a function of H)

3.10 SUGGESTED READING AND REFERENCES

Daniel, W. W. (1990) *Applied Non-parametric Statistics*, 2d ed. Boston: PWS-Kent.

Iman, R. L., and W. J. Conover (1981), Rank transformations as a bridge between parametric and non-parametric statistics, The American Statistician, 35, 124–129.

Siegel S. and Castellan N.J. (1988) *Non-parametric Statistics for the Behavioral Sciences* (2nd edition). New York: McGraw Hill.

References

Johnson, Morrell, and Schick (1992), Two-Sample Non-parametric Estimation and Confidence Intervals Under Truncation, *Biometrics*, 48, 1043-1056.

Leach, C. (1979). Introduction to statistics: A non-parametric approach for the social sciences. Chichester: John Wiley & Sons

Lehman, E. L. (1975). Non-parametric statistical methods based on ranks. San Francisco: Holden-Day.

Wampold BE & Drew CJ. (1990) Theory and application of statistics. New York: McGraw-Hill.