RESEARCH METHODS IN EDUCATION

Edited by

Gabriel O. Alegbeleye
Iyabo Mabawonku
Martins Fabunmi

Published by
The Faculty of Education
University of Ibadan,
Nigeria.

ISBN 978-076-502-6

Printed at Ibadan University Printery

List of Contributors

Abass, A. O.
Adedeji, S.
Adepoju, Tunji
Aderinoye, R. O.
Adesoji, Francis A.
Adewale, J. G.
Adeyemo, D. A.
Adu, E. O.
Ajayi, M. A.
Akinbote, O.
Akinwumiju, J. A.
Alegbeleye, G. O.
Ayodele-Bamisaiye, Oluremi
Babalola, J.
Babarinde, Kola
Eniola, M. S.
Fabunmi, Martins
Fadoju, A. O.
Hammed, Ayo
Jaiyeoba, Adebola O.
Moronkola, O. A.
Nwalo, Kenneth Ivo Ngozi
Nwazuoke, I. A.
Ogunleye, B. O.
Okediran, Abiodun
Onuka, A. O. U.
Osiki, J. O.
Osokoya, Israel Olu
Popoola, S. O.
Salami, S. O.

16

Non-parametric Analysis of Variance

J. G. Adewale

and

A. O. U. Onuka

Institute of Education,
University of Ibadan,
Ibadan – Nigeria.

Introduction

There are many non-parametric analyses available; some of them are chi-square, Mann Whitney U-test, Wilcoxon test, Kruskal-Wallis test, kendall-tau, friedman's test, etc. However, the focus of this chapter is the non-parametric analysis of variance. We may not understand the non-parametric analysis fully if an effort is not made to discuss its parametric component.

Parametric Statistics

Parametric test (statistic) is designed to represent a wide or large population, for instance, a country, or age group. Put differently, a parametric statistical test depends on certain assumptions about the population from which the sample is drawn. They are that:

1. The population scores are normally distributed (Kerlinger and Lee, 2000). This assumption of ***normality*** of distribution in the universe that contains in used in the research sample that produces the data being analysed must be plausible.
2. Homogeneity of variance is another assumption behind the use of parametric statistical testing of significance such as t-tests and F-distribution tests of significance.
3. There is the assumption of continuity and equal intervals of measures. This means that the measures used in the study must be continuous with equal intervals.
4. Yet another assumption is that there should be independence of observations in the study.

Therefore, parametric tests are based on the assumptions that the parameters or the characteristics of population containing the sample are known.

According to Cohen, Manion and Morrison (2003), these mean that:

(a) there is a normal curve of distribution of scores in the universe,

(b) the measures of scores in the population are continuous and possess equal intervals.

Gay and Airasian (2000) shared this view by stating that in order to use parametric statistical tests, the above enumerated conditions or assumptions must be met, if their results are to be valid. When one or two of these assumptions are not met, then the appropriate tests to use are the non-parametric statistical tests which are not dependent on these assumptions. This is so because the non-parametric makes no such assumptions about the shape of the distribution in the population of study. If the data are based on interval or ratio measures the parametric statistics are the obvious analysis to use in the test of significance, unless one or more of the above assumptions is/are violated. According to Babbie (1992) conditions for using parametric statistics are the assumptions that:

1. the independence of the two variables in the universe,
2. representativeness of samples selected from the population through probability sampling procedures,
3. the observed joint distribution of sample elements in terms of the variables.

Furthermore, Kerlinger (1973) believes that a parametric test is a product of statistical analysis that is made from research samples drawn from a population whose characteristics of normality, homogeneity of variance and continuity and equal intervals of its measure can and are assumed. Thus, it can be inferred, that only data, from samples whose universe can be assumed to possess homogeneity in its variance, possesses the propensity for continuity and equality of interval in its measure, lends itself to parametric analysis of variance. Otherwise, it is a non parametric analysis of variance that should be employed.

Results from parametric tests are more widely applicable because these aforesaid underlying assumptions about the population, from which the study sample was drawn. Whereas, it is more difficult to prove the generalizability of the results of the non-parametric test due to the absence of such assumptions about the population that produced the study sample. Cohen, Manion and Morrison (2003) explained that parametric tests are more powerful than non-parametric tests because they are not only derived from standardized scores but they enable the researcher to compare sub-populations with whole populations. This is due to the fact that it utilizes more powerful statistics such as means, standard deviation, t-tests, Pearson product moment, and factor analysis among others.

Non-parametric Statistics

Non-parametric statistical tests are applied when it is not feasible to make the assumptions underlying the use of the non – parametric statistics. Thus, non parametric tests are used only when it is possible to make only a few or no assumptions about the distribution of population from which the sample is drawn. Siegel in Kerlinger, (1973) defines a non-parametric statistical test as a test whose model does not specify conditions about the parameters of the population from which the sample is drawn.' Non-parametric statistics depend on ranking. Since parametric tests do not make assumptions about the wider population, a different type of statistics is used. These include modal scores, rankings, the chi-square statistic, and Spearman Brown correlation. They can be employed in very specific situations - like one class of students, one year group, one style of teaching, and one curriculum area and thus available for use by the teacher (Cohen, Manion and Morrison, 2003). According to Gay and Airasian (2000), once any one or two of the underlying assumptions for using parametric tests is absent in the population of which the investigation sample is a product or can be made, then the best statistic for testing significance in such study is the non-parametric statistic since the latter statistics do not make any assumption about the shape of the distribution (population that produces the sample).

Non-parametric tests are used when measures in a research are in the ordinal or nominal scale, when parametric assumptions have been violated, or when the nature of distribution is not known. A non-parametric statistic may not be able to perform tests on certain hypotheses which the parametric can. A larger sample size is needed in a non-parametric statistic to reach the same level of significance as the parametric.

Properties of Non-parametric Methods

The properties of non-parametric methods are as follows:

- Data can be tested against chance: e.g. coin-tossing - odds and evens, which are dichotomous data;
- Periodicity- data possesses the property of being periodic in nature;
- The distribution of data from different samples can be compared with each other or with a criterion group;
- A set of data is in rank order, when data can be ranked.

Advantages and Disadvantages of Non-parametric Statistics

The main merits of non-parametric statistics are that they do not impose the requirements of homogeneity and normality among others on the data as well as the fact that they are easier to calculate, which advantage the use of computer has eliminated. The demerits include the fact that they are less powerful in discovering differences when they exist. It is less detailed than the parametric statistics. They are also less specific on conclusions on a hypothesis testing.

Non-parametric Analysis of Variance

Kerlinger and Lee (2000) believe that non-parametric analysis of variance or methods depend on ranking. Abosede and Adesoye (2000) state that non-parametric statistics is used when the premise or form cannot be made i.e. when we cannot determine or assume any value of a parameter. Vaughan (2001) asserts that non-parametric statistics become useful when data type are in ordinal form or when one set of the data in consideration is in ordinal and the other(s) is/are in interval of ratio form(s) or yet if both or all sets of data are in interval/ratio but do not possess the requirement for using parametric statistics as earlier outlined. While its calculation is straight-forward it is limited by its inability to give estimates of means and to provide confidence intervals. It is equally identified with type II error, that is, accepting a hypothesis that should otherwise have been rejected. For a non-parametric statistic to attain the efficient in detecting the differences between means as does the parametric much larger sample. Data used in non-parametric analysis need not meet the same conditions as the parametric as earlier outlined. It needs only exhibit randomness and no underlying continuity. In many instances, we are not sure of the distribution pattern of the variables of interest in the population from the sample that was drawn. For example, income is not likely to be normally distributed in the population. Also the rate of occurrence of rare diseases is unlikely to be normally distributed in the population, the number of car accidents is also not normally distributed, and neither are very many other variables in which a researcher might be interested.

Sample size is another factor that often limits the applicability of tests based on the assumption that the sampling distribution is normal in the size of the sample of data available for the analysis (***sample size***; ***n***). However, we may assume that the sampling distribution is normal even if we are not sure that the distribution of the variable in the population is normal, as long as our sample is large enough (e.g., 100 or more observations). But, if our sample is very small we may use parametric tests only when we are sure that the variable is normally distributed. But there is no way to test this assumption if the sample is small, except we are certain of the distribution pattern of the variable of investigation within the population.

There are several types of non-parametric analysis. The following are some non-parametric analysis of variance among others: One-way analysis of variance otherwise known as ***the Kruskal-Wallis Test*** (Kerlinger and Lee, 2000). Abosede and Adesoye (2000) add that one way analysis of variance by ranks is for comparing three or more distributions. It can be used for both experimental and non experimental data. It is simple, but effective. Also for the two-way analysis of variance the Friedman Test is the non-parametric statistic that would be employed. This is used in situations when subjects are matched or the same subjects are observed more than once. It is a form of a rank-order analysis of variance.

The most common non-parametric analyses of variance are the Kruskal-Wallis Test for one-way analysis of variance when the data do not fall within the realm of the parametric and the Friedman Test for the way analysis of variance. Ideally measurement issue is usually discussed in statistics with respect to ***types of measurement*** or ***scale of measurement.*** The most common statistical techniques such as analysis of variance (and t-tests), regression, etc. assume that the underlying measurements are at least of ***interval*** implying that equally spaced intervals on the scale can be compared in a meaningful manner (e.g, K minus L is equal to M minus C). This assumption is sometimes not tenable, and the data rather represents a *rank* ordering of observations (ordinal) rather than precise measurements. According to Oyesiku (2000) the parametric statistics are used within the realm of inferential statistic in the management and social sciences because the assumption of normality in the distribution of variables in the population cannot be fixated, though non-parametric is less robust and less powerful than the parametric. Therefore, non-parametric is used mainly when the less sophisticated statistical analysis is requested, the test does not need the estimation of the population parameter. For Vaughan (2001) each of the non-parametric type has its parametric equivalent. For instance, the Spearman Correlation Coefficient is used in the place of Pearson Correlation, the Mann-Whitney test in place of the independent t-test, the Wilcoxon signed test instead of the paired test and the Kruskal-Wallis for the one-way ANOVA while the Chi-Square is used to measure association. These are so only when data set(s) do not meet the requirements for the use of parametric statistics. In cases where the data are in interval/ratio forms but do not qualify for the use of parametric statistics, such data are converted into ordinal data before applying the non-parametric equivalent of appropriate parametric statistics.

Non-parametric Methods

The need for statistical procedures that allow for processing data of low quality, from small samples, on variables about which nothing is known (concerning their distribution) is evident. Specifically, non-parametric methods were developed to be used in cases when the researcher knows nothing about the parameters of the variable of interest in the population (hence the name ***non-parametric***). In more technical terms, non-parametric methods do not rely on the estimation of parameters (such as the mean or the standard deviation of a universe) describing the distribution of the variable of interest in the population. Therefore, these methods are also sometimes (and more appropriately) called ***parameter-free*** methods or ***distribution-free*** methods.

Basically, there is at least one non-parametric equivalent for each parametric general type of test. In general, these tests fall into the following categories:

- Tests of differences between groups (independent samples)
- Tests of differences between variables (dependent samples)
- Tests of relationships between variables.

Differences between Independent Groups

When comparing two samples concerning their mean value for some variable(s) of interest, we should normally use the *t*-test (for independent samples); non-parametric alternatives for this test are the ***Wald-Wolfowitz runs test***, the ***Mann-Whitney U test***, and the ***Kolmogorov-Smirnov two-sample test***. In the case of multiple groups, one would ordinarily use analysis of variance and/or analysis of covariance (ANOVA/ANCOVA); the non-parametric equivalents to this method are the ***Kruskal-Wallis analysis of ranks*** **and the** ***Median test***. In this instance, the Kruskal-Wallis is of major interest to us as we will see later.

When a Particular Method should be Used

It is a difficult task to give simple advice concerning the use of non-parametric procedures. Each non-parametric procedure has its peculiar sensitivities and blind spots. For example, the Kolmogorov-Smirnov two-sample test is not only sensitive to differences in the location of distributions (for example, differences in means) but is also greatly affected by differences in their shapes. The Wilcoxon matched pairs test assumes that one can rank the data in the order of the magnitude of differences in matched observations in a meaningful manner. If this is not the case, one should rather use the Sign test.

Generally, when a particular result of an investigation is important (for example, does a very expensive and painful drug therapy help people get better?), it would always be advisable to run different non-parametric tests. Should discrepancies in the results occur contingent upon which test is used, one should try to understand why some tests give different results. On the other hand, non-parametric statistics are less statistically powerful (sensitive) than their parametric counterparts, and if it is important to detect even small effects (e.g., is this food additive harmful to people?) one should be very careful in the choice of a test statistic.

Large Data Sets and Non-parametric Methods

Non-parametric methods are most appropriate when the sample sizes are small. When the data set is large (e.g., $n > 100$) it often makes little sense to use non-parametric statistics at all. Whenever very large sample sizes are used in a study, the sample means will follow the normal distribution, though the respective characteristic is not normally distributed in the population, or is not measured very well. Thus, parametric methods, which are usually much more sensitive, (i.e., have better *statistical power*) are in almost all cases, appropriate for large samples. However, the tests of significance of many of the non-parametric statistics described here are based on the asymptotic (large sample) theory; therefore, meaningful tests can often not be performed if the sample sizes become too small.

Some Examples of Non-Parametric Analysis of Variance

One-Way Analysis of Variance

Let us assume that one is studying the differences in the non-objectivity of the management style of the members of three different State Basic Universal Education Boards (SUBEB's) in Nigeria, but find it difficult to administer an instrument on that purpose on them and thus requested an expert to judge to rank all the members of three Boards on the basis of a one-on-one interview with the members of the boards if membership of boards I, II and III are six, six and five respectively. The data from the interview are as presented in the table below:

Example 1

Ranks of seventeen members of three SUBEB's on differences in non-objectivity in management style.

BOARDS

	1	11	111
	12	11	4
	14	16	8
	16	8	3
	11	5	9
	14	4	1
	14	4	-
Sigma Ranks	81	47	25

The formula is

$$H = \frac{12 \Sigma Rj^2}{N(N+1)\ nj} - 3(N+1) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

Where N = total number of ranks

Nj = number of ranks in group j

Rj = sum of ranks in group j

Solution

$$\frac{12}{N(N+1)} = \frac{12}{17(17+1)}$$

$$= 0.03922$$

$$\frac{\Sigma Rj^2}{Nj} = \frac{(81^2)}{6} + \frac{(47^2)}{6} + \frac{(25^2)}{5}$$

$$= 1093.50 + 368.17 + 125.0$$

$$= 1586.67$$

$$3(N+1) = 3(17+1)$$

$$= 54$$

Now substituting into the equation (1)

$$H = \frac{12}{17(17+1)} \frac{\Sigma Rj^2}{nj} - 3(N+1)$$

$$= 0.03922 \times 1586.67 - 54$$

$$= 62.22 - 54$$

$$= 8.22$$

H is approximately distributed as X^2. The degree of freedom (df) is K – 1 where K is the number of columns or groups. Hence, since df = 3 – 1 = 2. Checking the $X^{2\ table}$ at 5 percent significant level (0.05), the critical value is 5.99. This value is less than the observed value (8.22), we can therefore conclude that there is a significant difference in objectivity in management style among the three SUBEB's boards.

Example 2

Three professionals were asked to examine thirty students (who have been identified to have some deviant behaviours) who were brought before them for counselling, the first being a psychologist, the second being a teacher and the third being a parent. The results of their findings are herewith presented:

1	2	3
44	23	11
47	29	16
49	34	18
74	20	15
81	28	14
87	27	14
84	21	09
52	23	10
67	30	11
74	20	15
659	**255**	**133**

$(Rj)^2$ = 434281 $(Rj)^2$ = 650250 $(Rj)^2$ = 176890

$$\frac{(Rj)^2}{nj} = \frac{434281}{10} = 43428.1 \qquad \frac{(Rj)^2}{nj} = \frac{65025}{10} = 6502.5$$

$$\frac{(Rj)^2}{nj} = \frac{176890}{10} = 17689$$

$$H = \frac{(12)}{N(n+1)} \quad \frac{\Sigma(Rj)^2}{nj} \quad - 3(n+1)$$

$$\frac{12}{n(n+1)} = \frac{12}{30(30+1)} = \frac{12}{(30)(31)} = \frac{12}{930}$$

$$\frac{\Sigma(Rj)^2}{nj} = 43428.1 + 6802.5 + 17689 = 67919.6$$

$$3(n + 1) = 3(30 + 1 = 3 \times 31 = 93$$

Substituting in the equation

$$= 0.0129 \times 67919.6 - 93$$

$$= 876.163 - 93$$

$$= 783.163$$

The degree of freedom (df) is 2. Checking the X^2 table at 5 percent significant level (0.05), the critical value is 5.99. Since this value is less than the observed value of 783.163, we can therefore conclude that there is a significant difference in the observations of the three professionals who counselled the deviant students.

Example 3

Three observers were asked to examine some sets of kindergarten pupils, eight pupils per an observer, but the last observer was only able to examine seven pupils with respect to their ability to pronounce the following words ***stamp your feet on Tuesday***. The results of their findings are presented in the following table.

1	2	3
17	16	4
14	14	3
16	12	1
11	10	-
20	08	2
24	06	5
18	04	4
40	02	3
160	**72**	**22**

$$\frac{(Rj)^2}{nj} = 25{,}600 \qquad 5184 \qquad 484$$

$$\frac{(Rj)^2}{nj} = \frac{25600}{8} \qquad \frac{5184}{8} \qquad \frac{484}{7}$$

$$= 3{,}200 \qquad 648 \qquad 69.14$$

$$\frac{12}{n(n+1)} = \frac{12}{23(23+1)} = \frac{12}{23x24} = \frac{1}{46}$$

$$\frac{\Sigma Rj^2}{nj} = 3{,}200 + 648 + 69.14 = 3917.14$$

$$3(n + 1) = 3(23 + 1) = 3 \times 24 = 72$$

$$H = \frac{12}{N(n+1)} = \frac{\Sigma Rj^2}{nj} - 3(n + 1)$$

$$= \frac{3917.14}{46} - 72$$

$$= 85.155 - 72$$

$$= 13.155$$

$$= 13.16$$

The degree of freedom (df) is 2. Checking the $X^{2\ \text{table}}$ at 5 percent significant level (0.05), the critical value is 5.99. This value is less than the observed value of 13.16, we can therefore conclude that there is a significant difference in the examinations of the three teachers.

Two-Way Analysis of Variance: The Friedman Test

In the previous section, we discussed non-parametric one-way analysis of variance, there is also the non-parametric two-way analysis of variance. Just as Kruskal and Wallis worked on the non-parametric one-way analysis of variance, Fredman worked on the non-parametric two-way analysis of variance. This is applicable in situations in which subjects are matched or the same subjects are observed more than once, a form of rank-order analysis of variance. When it is not feasible to use the Kruskal and Wallis formula in a non-parametric analysis, because the results are not amenable to the one – way analysis, then the Friedman Test becomes the obvious analytical tool to employ. The Friedman test does not fulfil the assumptions of the parametric method. The method is not suitable for data where there is more than one observation in each cell of the two way tables. It assumes that there are no ties in the data for each group but will be little affected by a few ties. It is assumed that the Friedman's analysis with to groups is equivalent to an extension of the sign test rather than the Wilcoxon test. The ranking in each row is assumed to be separate and independent from the ranking in all other rows. It does not require a complete ranking of all scores in matrix as found in the Kruskal Wallis test.

The formula is as follows:

$$H = \frac{12}{Kn(n+1)} \Sigma Rj^2 - 3K(n-1)$$

Where

Chi-square, ranks;

k is the no of rankings;

n the no of objects being ranked;

R the sum of the ranks in each column; and

Sigma R square is the sum of these squared sums of the ranks.

Examples (Hypothetical)

The following table shows some data from two achievement tests to compare the overall grades scored by four students in a final year exam. The values for the four students are ranked in the various subjects taught. At 0.01 level of significance, discover the poorest performing student.

Subject	**A**	**B**	**C**	**D**
Maths	308	132	454	64
English	102	526	0	28
Physis	182	134	96	30
Biology	268	324	264	90
H/Econs	166	228	134	34
Agric	332	296	458	6
Geography	198	350	200	90
History	28	274	16	24
Mean	**198**	**283**	**203**	**45.7**
Standard Deviation SD	**103**	**127**	**179**	**31.6**

Ranks of the Data in Table 1.2a Above.

Subject	A	B	C	D
Maths	3	2	4	1
English	3	4	1	2
Physis	4	3	2	1
Biology	3	4	2	1
H/Econs	3	4	2	1
Agric	3	2	4	1
Geography	2	4	3	1
History	3	4	1	1
Total Ri	24	27	19	10
Mean Rank	**3.00**	**3.38**	**2.38**	**1.25**

The analysis proceeds in a similar way to the Kruskal Wallis test.

Using the equation given on statistics H, we calculate H as:

$$H = \frac{12}{8 \times 4 \times 5} (24^2 + 27^2 + 19^2 + 10^2) - 3 \times 8 \times 5$$

$$H = 12.45$$

A reference to the table of chi-square for 3 df (number of treatment – 1) shows that the observed X^2 value of 12.45 is significant at 5 percent level (0.05 = 7.82). We can therefore conclude that the achievement ranks of the four students in different subjects is significant at a 0.05 level of significance.

References

Abosede, A. J. and Adesoye, B. (2000) Hypothesis Formulation and Testing in Odugbemi, O. O. and Oyesiku, O. K. (eds.) Research Methods in the Social and Management Sciences. Ago-Iwoye: CESAP, Ogun State University. 156-173

Babbie, Earl (1992) The Practice of Social Research. 6th edition. Belmont, California: Wadsworth Publishing Company.

Cohen, L.; Manion, L. and Morrison, K. (2003) Research Methods in Education 5th edition. A reprint. London: Routledge Falmer.

Gay, L. R. and Airasian, P. (2000) Educational Research – Competencies for Analysis and Application. 6th edition. New Jersey: Prentice – Hall Inc.

Kerlinger, F. N.(1973) Foundations of Behavioral Research. 2nd Edition. New York: Holt, Rinehart and Winston, Inc.

Kerlinger, F. N. and Lee, H (2000) Foundations of Behavioral Research. 4th Edition. USA: Wadsworth.

Oyesiku, O. K. (2000) Statistical and Data Analysis in Research in Odugbemi, O. O. and Oyesiku, O. K. (eds.) Research Methods in the Social and Management Sciences. Ago-Iwoye: CESAP, Ogun State University. 334-379.

Siegel, S. and Castellan, N. J. (1988) Non-parametric Statistics. 2nd Edition. New York: McGraw-Hill.

Vaughan, L. (2001) Statistical Methods for the Information professional –A Practical, Painless Approach to Understanding, using, and Interpreting Statistics. *American Society for Information Science and Technology (ASIST) Monograph Series*