Prestige Journal of Education, Vol. 6, No. 1, June 2023
A publication of the Association for the Promotion of Innovation in Education (APIE)
ISSN: 2645-324X (Print)
ISSN: 2645-3223 (Online)

The Application of the Nominal Scale of Measurement in Research Data Analysis

Delight Omoji Idika, Ph.D
delightoidika@yahoo.com
Institute of Education
University of Calabar, Calabar

[1]**Valentine Joseph Owan**
owanvalentine@gmail.com

[1]**Victor Ubugha Agama**
vygamafes@gmail.com
[1]*Department of Educational Foundations*
University of Calabar, Calabar

Abstract
Appropriate measurement scales are fundamental in data analysis, allowing researchers to categorise, select appropriate statistical methods, and analyse and interpret their data accurately. The nominal scale is one such measurement scale in behavioural sciences, which is crucial in organising data into distinct categories. This paper provides an overview of the nominal measurement scale in research data analysis. It explains the characteristics and role of the nominal scale in organising data into distinct categories. The paper discusses methods of collecting nominal scale data, including surveys and observations. It explores the use of the nominal scale in descriptive (such as frequency counts, measures of dispersion and central tendencies), and inferential statistics (such as point biserial correlation, independent t-test, analysis of variance, logistic regression, discriminant analysis, differential item functioning, chi-square test of independence, Kruskal-Walli's test, and Mann-Whitney U Test). Each technique is explained with assumptions and application areas. In conclusion, the paper emphasises the significance of the nominal scale in data analysis and its contribution to various statistical techniques. It serves as a comprehensive guide for researchers and practitioners looking to understand and utilise the nominal measurement scale in their data analysis.
Keywords: ANOVA, categorical, data, logistic, regression

Introduction

Measurement scales are essential in data analysis because they serve as the hub for statistical analysis. Statistical analysis, be it descriptive or inferential, relies on measurement scales. Therefore, understanding the different scales of measurement and their applications is the starting point of quantitative coding and data analysis. According to Stevens (1946), statistical manipulations that can legitimately be applied to empirical data depend upon the type of scale against which the data are ordered. Since measurement is defined as assigning numerical values to things or occurrences following established norms (Stevens, 1946), the fact that numbers may be allocated according to various criteria leads to various scales and measuring methods. A new challenge arises when it is necessary to specify which rules govern number assignment, which scale features (or group structure) are produced, and which statistical procedures are relevant to measurements conducted with each kind of scale.

Due to the similarity between what can be done with things and the features of numerical series, scales are, first and foremost, feasible. For example, empirical operations are used to determine equality (classifying), rank order and when differences and ratios between objects' characteristics are equal. It is possible to utilise the number series as a model to depict features of the empirical world because of its similarity to certain empirical procedures. The nature of the fundamental empirical processes must be considered to obtain a certain scale. The nature of the entity being scaled and the techniques one chooses usually restrict these operations but once picked, the operations decide whether one of the scales will emerge. According to the seminal work of Stevens (1946), there are four scales of measurement. These include the nominal, ordinal, interval and ratio scales in ascending order of their complexity and sophistication. The nominal scale categorises events, constructs, observations and phenomena based on equality. The ordinal scales quantify events or observations based on the magnitude (greater or less). The interval scale quantifies observations by determining the equality of intervals or differences. The ratio scale is based on the quantification of observations by determining the equality of ratios. Measurement scales and statistical approaches for data analysis are both dependent on the kind of data being gathered (Stevens, 1968). This paper discusses the nominal measurement scale and its application in data analysis.

Nominal scale of measurement

The nominal measurement scale is one of the four scales introduced by Stevens. It uses numbers qualitatively to classify events or observations by grouping them according to a common or shared qualitative attribute. Unlike other scales that are more quantitative in using numbers, the nominal scale makes arbitrary use of numbers. According to Howitt and Cramer (2011), "in nominal measurement, the variable consists of named categories that have no mathematical properties" (p.46). By implication, this measurement scale labels the things being measured. In education, psychology, social sciences and behavioural sciences, quantifying certain events, constructs, phenomena or observations cannot be meaningful without proper classification. This gives the nominal scale of measurement an important place in statistical data analysis. Although one can artificially classify events or observations into different rational groups, some variables are naturally dichotomised and do not lend themselves to rational categorisation.

For instance, one can artificially assign numbers to all the contestants in a marathon race. These numbers are to identify the contestant. The numbers are not associated with the result of the race or the person's characteristics. One can also classify the countries in the world according to the continents. When this is done, any country from Africa (such as Nigeria, Ghana, Morocco, Egypt and so on) can be placed in Group 1, countries from Asia in Group 2, countries from Europe in Group 3 and so on. It must be noted that the assignment of 1, 2, 3 and so on, for different continents does not imply that one group is richer or greater than the others or the other way around. Also, students in a classroom can be grouped according to their skin colour, such as black, brown, white, chocolate or yellow. This means that students of the same colour can be grouped. Although colour names are linguistic identifiers, one can use letters of the alphabet or numbers to represent each group if desired. Whether one uses the numbers 1, 2, 3, 4 and 5 to represent black, brown, white, chocolate and yellow skin colours, respectively, does not make any practical difference. In contrast, variables such as sex, place of origin, marital

ISSN: 2645-324X (Print)
ISSN: 2645-3223 (Online)

status, genotype, blood type, race, eye colour and others are based on natural classification.

The application of the nominal scale of measurement in data analysis

This section discusses the application of the nominal scale of measurement in data analysis. Since data collection precedes data analysis, this section also discusses the methods of collecting nominal scale data and the descriptive and inferential statistical analysis that can be applied.

Methods of collecting nominal scale data

One must use a question-type survey to collect data using the nominal scale. Such questions could be of several kinds. Today's different online survey tools have many options for nominal scale. Nominal scale data can be collected through open-ended questions (questions that permit the respondent to answer freely) and multiple responses question (that allows the respondent to select more than one option as their choice of an answer). It may include questions followed by a blank space to answer (For example, what is your name?). It may include a question and a list of options to select from (for example, what is your natural hair colour? Then a list of hair colours is presented). Nominal data can also be collected using multiple-choice items to determine the number of students that picked different options. Checkboxes can also be used to obtain nominal data by providing a list and allowing respondents to check the one(s) that applies to them (for example, sex [male or female], marital status [single or married], or location [urban, suburban and rural] and so on). When the Likert scale is used, nominal data can be derived by counting the number of respondents that strongly agreed, agreed, disagreed or strongly disagreed with statements. In this case, numerical weightings of response options (such as 4, 3, 2, 1 or 1, 2, 3, 4, depending on item wording) are unnecessary.

The nominal scale of measurement in descriptive statistical analysis

Because nominal scales are often qualitative, the issue of calculation and analysis is often confusing. The numbers assigned to the attributes have no numerical value. This means that no arithmetic computation can be performed on data obtained by nominal scale. However, many statistical or mathematical operations are possible with the nominal scale. One cannot perform basic arithmetic operations such as addition, subtraction, multiplication, and division on nominal-scale data. One cannot also check for inequalities such as less than, equal to, or greater than by mere group membership. However, one can count the number of elements in each category. If the observations used for the categorisation were derived from the same population, one could perform a simple percentage analysis to determine the proportion of elements of the population that belong to different categories. Statistically, a measure of central tendency (mode) applies to the nominal scale but not others, such as mean and median. That is, the category with the most elements can be indicated. For example, the skin colour that characterises most students in a class can be told.

According to Howitt and Cramer (2011), the only statistical analyses possible with the nominal scale are those based on frequency data. For instance, after collecting data based on the religious affiliation of all the people in an institute, the researcher can categorise the responders into separate groups. After segmenting the responders, he can quantify the percentage of each group against the total number of participants. On the other hand, the median or mean cannot be calculated because they would make no sense. The qualities of the attribute can be put in any order since ranking the variables are meaningless. Graphical interpretation using pie charts and bar charts are two techniques for analysing nominal data. One can use a pie chart to represent the degree to which elements in a population can be found across categories. When using a bar chart, the height of each bar can represent the frequency of responses in each category. All the statistical analyses discussed above are descriptive, but some inferential statistical analyses are also possible with nominal scale data. When it comes to inferential data analysis, care must be taken to ensure that other assumptions for using them are met.

The nominal scale of measurement in inferential data analysis
Inferential statistical analyses possible with the nominal scale of measurement include parametric tests such as point biserial correlation, independent t-test, Analysis of variance, binary logistic regression analysis, multinomial logistic regression, discriminant analysis, differential item functioning and others. Statistical methods that can be used to compare two or more groups of things may apply to the nominal level of measurement. However, non-parametric tests such as the Chi-square test of independence, Mann-Whitney U test, Kruskal-Wallis' test, Friedman test, and Mood's Median test can also use the nominal scale of measurement. The application of the nominal measurement scale for each of the statistical methods mentioned above is discussed below.

3. Point biserial correlation

A point-biserial correlation is used to measure the strength and direction of the association between one continuous variable and one dichotomous variable. For example, one could use a point-biserial correlation to determine whether there is an association between salaries, measured in US dollars, and gender (that is, the continuous variable would be "salary" while the dichotomous variable would be "gender", which has two categories: "males" and "females"). It is only appropriate to use a point-biserial correlation if the data "passes" this assumption. The gender variable in the example is nominal but used with another continuous variable. This means that the data gotten from each participant must be partitioned into two nominal or ordinal groups.

4. Independent t-test analysis

Independent sample t-test is a statistical technique used to analyse the mean comparison of two independent groups. In an independent samples t-test, when two samples are taken from the same population, the mean of the two samples may be identical. However, when samples are taken from two different populations, then the mean of the sample may differ. In this case, it is used to compare the means of two populations and to tell whether they are similar. The independent test assumes that the dependent variable is normally distributed; the variance of the two groups is the same as the dependent variable; the two samples are independent of each other; samples are drawn from the population at random; all observations must be independent of each other; the dependent variables must be measured on an interval or ratio scale.

It must be noted that the nominal scale of measurement applies to the independent t-test through the independent variable that represents two groups. Although the independent variable in an independent t-test analysis may sometimes be ordinal, the dependent variable must always be continuous. For example, one can compare students' performance from rural and urban schools

ISSN: 2645-324X (Print)
ISSN: 2645-3223 (Online)

(nominal variable) in mathematics achievement tests. One can also test for gender differences in completing a race (where race, the dependent variable, is continuously measured with time). Students' achievement in an English language test can be compared among students with favourable or unfavourable attitudes towards learning (attitude, in this case, represents two nominal groups).

5. Analysis of variance

Analysis of variance (ANOVA) is used to determine whether there are any statistically significant differences between the means of three or more independent (unrelated) groups based on a continuous dependent variable. The independent variable with three or more levels is usually nominal or ordinal (categorical), whereas the dependent variable is continuous. Although one tends only to see it used when there is a minimum of three levels in the independent variable, it can also handle data from two nominal or ordinal groups. According to Chanal et al. (2021), three assumptions must be satisfied with ANOVA F-Test: samples are independent, from a normally distributed population, and standard deviations of the groups are all equal (homoscedasticity). There are different kinds of ANOVA: one-way, with just a single factor and two- or multiway, with two or more factors, and main- and interaction-effects models (David, 2017).

For example, one could use a one-way ANOVA to understand whether an international exam performance differed based on students' nationality. The nationality of students in this example is nominal. Two-way ANOVA can determine whether students' interests in physics differ based on the school location (rural, suburban and urban) and gender (male and female). In this example, one can tell whether there are differences between males and females in rural, suburban and urban locations based on their interest in Physics. It can also be established whether students' interest in physics differs among males or females across the three locations. The nominal measurement scale can be applied in almost all ANOVA scenarios depending on the study's objective. The ANOVA is an omnibus test statistic and cannot tell which specific groups were statistically significantly different from each other; it only tells that at least two groups were different. Determining which groups differ is important since one may have three, four, five or more groups in one’s study design. This can be done using a post hoc test.

6. Binary and multinomial logistic regression analysis

Logistic regression is generally a statistical data analytic procedure used to understand the relationship between the dependent variable and one or more independent variables by estimating probabilities using a logistic regression equation. This type of analysis can help to predict the likelihood of an event happening or a choice being made. For instance, logistic regression can be used to determine whether a group of students are likely to pass or fail an exam or perform highly, averagely or poorly based on certain variables such as gender, school location, anxiety levels, attitude, gender, among others. The logistic model is so powerful that it can handle data at all measurement scales relatively or simultaneously as long as its assumptions are met.

The Binary logistic model (or logit model) is used to model the probability of a certain class or event existing, such as pass/fail, win/lose, alive/dead or healthy/sick (for binary regression). These possible events occurrences are nominal but can be ordinal with two levels too (for example, high/low). Binary logistic regression uses a logistic function to model a binary dependent variable. Mathematically, a binary logistic model has a dependent variable with two possible values, such as pass/fail, represented by an indicator variable, where the two values are labelled "0" and "1". The numbers used here are for identification, classification or grouping purposes and do not possess magnitude (nominal scale). In a binary logistic model, the log-odds (the logarithm of the odds) for the value labelled "1" is a linear combination of one or more independent variables ("predictors"); the independent variables can each be a binary variable (two classes, coded by an indicator variable) or a continuous variable (Tolles & Meurer, 2016).

On the other hand, multinomial logistic regression is a classification method that generalises logistic regression to multiclass problems (that is, with more than two possible discrete outcomes), according to Greene (2012). It is a model used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable given a set of independent variables (which may be real-valued, binary-valued, categorical-valued, and so on). Multinomial logistic regression is a simple extension of binary logistic regression that allows for more than two categories of the dependent or outcome variable. Like binary logistic regression, multinomial logistic regression uses maximum likelihood estimation to evaluate the probability of categorical membership. The categorical membership in question is nominal (or equivalently categorical, meaning that it falls into any one of a set of categories that cannot be ordered in any meaningful way) and for which there are more than two categories.

7. Discriminant analysis

Discriminant analysis is a method researchers use to examine research data where the criterion or dependent variable is categorical (nominal or ordinal) and the predictor or independent variable is of the interval type. The phrase categorical variable refers to the fact that the dependent variable is classified into multiple groups. For example, the categorical dependent variable may be three computer brands: A, B, and C. The goal of discriminant analysis is to create discriminant functions, a linear combination of independent variables that perfectly discriminate between the categories of the dependent variable. It allows the researcher to see whether there are any significant differences between the groups in terms of the predictor factors. It also assesses the classification's correctness.

The number of categories held by the dependent variable characterises discriminant analysis. Everything is assumed up to infinity in statistics; therefore, when the dependent variable contains two categories, the type utilised is two-group discriminant analysis. Multiple discriminant analysis is performed when the dependent variable includes three or more categories. Several situations demonstrate why discriminant analysis is appropriate. It may be used to determine if heavy, medium and light soft drink consumers consume varying amounts of frozen meals. In the discipline of psychology, it may be used to distinguish between price-sensitive and non-price-sensitive grocery customers based on psychological qualities or features. It may be used in business to understand the traits or features of a consumer with store loyalty versus a client without. The general idea here is that the nominal measurement scale can be applied to discriminant analysis if the variables so warrant.

ISSN: 2645-324X (Print)
ISSN: 2645-3223 (Online)

8. Differential item functioning

In Item-Response Theory (IRT), Differential item functioning (DIF) is a statistical characteristic of an item that shows the extent to which the item might be measuring different abilities for members of separate subgroups. DIF refers to differences in the functioning of items across groups, often demographic, which are matched on the latent trait or, more generally, the attribute being measured by the items or test (Camilli, 2006; Holland & Wainer, 1993). Average item scores for subgroups having the same overall score on the test are compared to determine whether the item is measuring in essentially the same way for all subgroups. An item does not display DIF if people from different groups have a different probability of giving a certain response; it displays DIF if and only if people from different groups with the same underlying true ability have a different probability of giving a certain response. Common procedures for assessing DIF are Mantel-Haenszel, item response theory (IRT) based methods, and logistic regression (Zumbo, 2007). The nominal scale of measurement has application to DIF because it is often used to discriminate the functioning of items based on nominal variables such as gender, school location, and others.

9. Chi-square test of independence

The Chi-Square test of independence is used to determine if there is a significant relationship between two nominal (categorical) variables. The frequency of each category for one nominal variable is compared across the categories of the second nominal variable. The data can be displayed in a contingency table where each row represents a category for one variable, and each column represents a category for the other variable. For example, if a researcher wants to examine the relationship between gender (male vs. female) and empathy (high vs. low) the chi-square test of independence can be used to examine this relationship.

Before using the Chi-square, the following assumptions must be considered: the data in the cells should be frequencies or counts of cases rather than percentages or some other transformation of the data; the levels (or categories) of the variables are mutually exclusive (that is, a particular subject fits into one and only one level of each of the variables); each subject may contribute data to one and only one cell in the χ^2; the study groups must be independent (this means that a different test must be used if the two groups are related); there are two variables, and both are measured as categories, usually at the nominal level (however, data may be ordinal data; interval or ratio data that have been collapsed into ordinal categories may also be used).

10. Kruskal Wallis test

The Kruskal–Wallis test by ranks, Kruskal–Wallis H test or one-way ANOVA on ranks is a non-parametric method for testing whether samples originate from the same distribution (Kruskal & Wallis, 1952; Corder & Foreman, 2009). It compares two or more independent samples of equal or different sample sizes. It extends the Mann–Whitney U test, which compares only two groups. The parametric equivalent of the Kruskal–Wallis test is the one-way analysis of variance (ANOVA). A significant Kruskal–Wallis test indicates that at least one sample stochastically dominates one other sample. The test does not identify where this stochastic dominance occurs or for how many pairs of groups of stochastic dominance obtains. For analysing the specific sample pairs for stochastic dominance, Dunn's test (Dunn, 1964), pairwise Mann–Whitney tests with Bonferroni correction, or the more powerful but less wellknown Conover–Iman test (Conover & Iman, 1979) are sometimes used. Just like in ANOVA, the Kruskal-Wallis test can be applied to data collected at the nominal scale of measurement. The Kruskal-Wallis test requires the independent variable to be in nominal or ordinal categories, whereas the dependent variable can be ordinal or continuous.

11. Mann-Whitney U Test

The Mann-Whitney U test compares differences between two independent groups when the dependent variable is ordinal or continuous but not normally distributed. Unlike the independent-samples t-test, the Mann-Whitney U test allows for drawing of different conclusions about data depending on the assumptions made about the data's distribution. These conclusions can range from simply stating whether the two populations differ, to determining if there are differences in medians between groups. The test assumes that the dependent variable is measured at the ordinal or continuous level; the independent variable consists of two categorical independent groups that are nominal or ordinal; observations are independent (meaning that there is no relationship between the observations in each group or between the groups themselves); it can be used when two variables are not normally distributed. Just like in an independent t-test, students' performance in Chemistry can be compared across two nominal or ordinal variables such as gender, school location, anxiety level (high or low) and so on.

Conclusion

This paper discussed the concept of scales of measurement with a specific focus on the nominal scale. The paper highlights the usefulness of the nominal scale in statistical data analysis. The methods of collecting nominal data are discussed, as well as an overview of some descriptive and inferential statistical techniques that can utilise the nominal measurement scale in data analysis. This paper concludes that the nominal scale of measurement is very important in quantitative data analysis. It can be used by almost all statistical techniques requiring categorical data. Apart from the logistic regression analysis that splits data into purely nominal categorical grouping, almost all other statistical methods that require categorical variables appear to work well with both the nominal and ordinal scales of measurement. In summary, it is pertinent to reiterate that using numbers on a nominal scale is for classification and possesses no reasonable mathematical or statistical meaning. Most of the statistical techniques applied to nominal data are often used to compare two or more groups based on a measured trait that may take any form from the ordinal to ratio scales (either as independent or dependent variables).

Recommendations

The following are the recommendations made:

1. Researchers should ensure proper understanding and implementation of the nominal scale of measurement when collecting and analysing data, as it provides a robust framework for categorising and interpreting categorical data accurately.
2. It is recommended to utilise various statistical techniques tailored for nominal scale data analysis, such as point biserial correlation, chi-square test of independence, and logistic regression, among others, to extract meaningful insights and draw reliable conclusions.

ISSN: 2645-324X (Print)
ISSN: 2645-3223 (Online)

3. Future studies should focus on exploring and developing advanced methods for analysing nominal scale data, such as incorporating machine learning algorithms or exploring more complex multivariate techniques, to enhance the depth and accuracy of data analysis.
4. Researchers should pay attention to the validity and reliability of the data collected on a nominal scale, ensuring appropriate measures are in place to minimise bias and errors during data collection and analysis, ultimately enhancing the credibility and robustness of research findings.

References

Camilli, G. (2006). Test fairness. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 220–256). Westport, CT: American Council on Education.

Chanal, D., Yousfi Steiner, N., Petrone, R., Chamagne, D., & Péra, M.-C. B. T.-R. M. (2021). *Online Diagnosis of PEM fuel Cell by Fuzzy C-means Clustering*. Amsterdam: Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-12-819723-3.00099-8

Conover, W. J., & Iman, R. L. (1979). *On Multiple-Comparisons Procedures.* New Mexico: Los Alamos Scientific Laboratory. https://doi.org/10.2172/6057803

Corder, G. W., & Foreman, D. I. (2009). *Nonparametric Statistics for Non-Statisticians*. Hoboken: John Wiley & Sons. https://doi.org/10.1002/9781118165881

David, V. (2017). 3 - Structure as groups of objects/variables. In V. David (Ed.), *Data Treatment in Environmental Sciences* (pp. 35–55). Amsterdam: Elsevier. https://doi.org/https://doi.org/10.1016/B978-1-78548-239-7.50003-6

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*, *6*(3), 241–252. https://doi.org/10.1080/00401706.1964.10490181

Greene, W. H. (2012). *Econometric Analysis* (7th ed.). Boston: Pearson Education.

Holland, P. W., & Wainer, H. (1993). *Differential Item Functioning.* Hillsdale, NJ: Lawrence Erlbaum.

Howitt, D., & Cramer, D. (2011). *Introduction to Research Methods in Psychology* (3^{rd} ed). Edinburgh: Pearson Education Ltd.

Kruskal, W., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441

Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, *103*(2684), 677-680. https://doi.org/10.1126/science.103.2684.677

Stevens, S. S. (1968). Measurement, statistics, and the schematic view. *Science*, *161*(3844), 849-856. https://doi.org/10.1126/science.161.3844.849

Tolles, J., & Meurer, W. J. (2016). Logistic regression relating patient characteristics to outcomes. *JAMA, 316(5), 533–4.* https://doi.org/10.1001/jama.2016.7653

Zumbo, B. D. (2007). Three generations of differential item functioning (DIF) analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly, 4,* 223–233. https://doi.org/10.1080/15434300701375832