Scholars Journal of Physics, Mathematics and Statistics

OPEN ACCESS

Abbreviated Key Title: Sch J Phys Math Stat
ISSN 2393-8056 (Print) | ISSN 2393-8064 (Online)
Journal homepage: https://saspublishers.com

Effect Size for Nested Design

Marcel Tochukwu Obinna[1*], Uchenna Petronilla Ogoke[1]

[1]Department of Mathematics and Statistics, University of Port Harcourt, Port Harcourt, Nigeria

DOI: 10.36347/sjpms.2021.v08i09.003 | **Received:** 02.10.2021 | **Accepted:** 08.11.2021 | **Published:** 29.11.2021

***Corresponding author:** Marcel Tochukwu Obinna

Abstract **Original Research Article**

Effect size is an efficient way of calculating/estimating the level of variation in any data set. It has been proven to give more information than hypothesis testing and p-value. It also enhances the information provided by hypothesis testing. Hence, several renowned scholars and organisations have recommended this technique as a pertinent analysis for meticulous discovery and conclusions on issues of level of significance/impact of factors. Effect size analysis is becoming ubiquitous with a widespread availability of both methods and software used for its estimation. Despite this growth, their use in nested design has been limited. This could be greatly attributed to the lack of non-ambiguous methods for estimating effect size in Nested Design. In this work, we will to extend Cohens (1988) formula to Nested Design. Providing novel and well-defined formulae that can be used to produce estimate for effect size in Nested Design for both fixed and random factors.

Keywords: Effect Size, NEDPY, Cohen's f for Nested Design, Nested Design, Effect size for fixed factors, Effect size for random factors.

1.0 INTRODUCTION

Conventionally, effect size is defined as the difference between ANOVA model parameters. It is also known as Cohen's f. It reflects the overall degree of heterogeneity among effect parameters (and population mean). Hypothesis has been discredited for its dichotomous nature. This inherent dichotomy is dissatisfying to researchers, who frequently use the null hypothesis as a statement of no effects while in fact they are more interested in knowing how big an effect is rather than whether it is zero or not. Even when the test (hypothesis test) identifies the presence of a significant factor/difference, it leaves the researcher uncertain as to the nature of this factor/difference. P-value does not necessarily solve this problem as p-value could be misleading as it is greatly affected by sample size. Estimating the effect size evidently resolves this quagmire as it clearly gives the researcher more insight about the nature and size of the impact caused by this factor thus, he/she could draw more accurate inferences.

Nested design is a type of experimental design used to analyse data set with nested factor(s). An experiment is said to have a nested factor if the levels of one factor (nested factor) occurs uniquely with only one level of the other factor. The application of Nested Design in real life is becoming enormous, ranging from application in modern science in areas such as medicine, ecology, statistics, horticulture, industrial technology, psychology to non –science fields like education, sociology etc. for example, to study the efficiency of hospitals on a particular medication. In this case, each patient occurs/interacts with a unique doctor and each doctor occurs/interacts with unique hospital. Hence, patients are nested within doctors and doctors are nested within a hospital. Using the common cross model will imply all the doctors should treat each patient and all the doctors will have to work in all the hospitals.

A great deal of emphasis has been made on the need to compute, present, and discuss effect size statistics as a routine part of any empirical report (American Psychological Association, 2001; Wilkinson & the Task Force on Statistical Inference, 1999). Several scholars have produced well–defined methods for determining effect size in other areas of experimental design and there have been lack of a general and conscience methods for such method for effect size estimation when dealing with data with nested structure. In this article, we hope to provide well-defined formulae for estimating effect size in Nested Design.

A predominant reason for running a research/experiment is to investigate whether or not a certain factor/treatment has an effect on a situation or to find out if there exists a difference between two or more different occurrences. A further scrutiny into this question will be to find out the size of the effect/difference (if it exists), thus necessitate effect size. The estimation of the effect size will determine if the effect size is small, medium or large.

Significance testing (hypothesis testing) has for a long time been used for research. It has permeated into virtually every area of research and has been customised by several scholars to handle data with different structures. Despite this "success", statistical hypothesis testing has received an enormous amount of criticism, and for a rather long time for the misleading nature of this approach. Johnson 1999 gave a detailed chronological account of some reputable critics, which plumply criticised the apt use of significance testing. Despite these critics, significance testing are still being used not minding the dangers of its misleading attributes. Johnson (1999) , Steiger & Fouladi (1997) gave some reasons why it is still being used which are; they appear to be objective and exact, they are readily available and easily invoked in many commercial statistic packages, everyone else seems to use them, students, statisticians, and scientists are taught to use them and some journal editors and thesis supervisors demand them.

Aside the above constrain, significance testing says nothing about the second most likely inquiry which is the effect/different size. Hypothesis testing gives a dichotomous conclusion and gives no measure or estimate to the effect when identified as significant. All these scholars suggest effect size estimation among other methods as a solution to this problem.

Effect size has existed as far back as 1800's. Huberty (2002) effectively reduced the perplexity associated with the history of effect size by dividing it into three groups for single output data, they are: relationship indices, group difference indices and group overlap indices. The relationship indices which is the oldest of the three is a form of indexing that estimates the relationship/correlation between two sets of variable - causation and output variable in an analysis unit. Francis Galton (1822 – 1912) first developed it as correlation. It was further developed into the coefficient of correlation by Francis Y. Edge (1892) worth and from then has been improved by several scholars, who made it gain application in ANOVA. It is popularly denoted with the symbol η or η^2.

Group overlapping indices – this effect size estimation is made based on the explicit measurement of the overlapping of two sets of data. It was first introduced by John W. Tilton (1891-1980) and was subsequently developed by Dunnette (1966), Alf and Abraham (1968) amongst others.

Huberty's third classification, which is apparently the most intuitive and common effect size index, is the group difference indices. Jacob Cohen (1923-1998) proposed this index. It is considered as the ratio of the variation among the group means to the average variation among subjects within each group as measured by their standard deviations. It is denoted by ***d*** for a data set with two groups and ***f*** for a data set with more than two groups. The values of ***f*** can range from zero upwards. Cohen (1988, 285-287) gives the following interpretation for ***f*** values. ***f*** = 0.10 is a *small* effect, ***f*** = 0.25 is a *medium* effect, ***f*** = 0.40 is a *large* effect. The formula later proposed in this article is under this subgroup and follows the same guideline as the Cohen's ***f*** estimate.

Effect size in Nested Design has been treated by Spyros Konstantopoulos (2008) who defined it as the ratio of the range of the effect parameter and the total standard deviation in the population. His definition was limited to only cases where the first factor has two levels i.e. control group and treatment group, because it was specifically designed for studying intervention program in the educational sector. Nianbo Dong (2014) also postulated formulae to calculate minimum detectable effect size (MDES) Nested Design, but these formulae were restricted to continuous moderator variable at the first or second level in a two-stage nested design with random factors. The method developed in this work can be used effectively for any variable type and different amount of level.

2.0 METHODOLOGY

This methods were constructed following Cohen's' 1988 guideline and follow its benchmark.

f = 0.10 is a *small* effect; ***f*** = 0.25 is a *medium* effect; ***f*** = 0.40 is a *large* effect.

For simplicity, we would consider the two-stage nested design. Note that the formulae formulated in this work can be analogously applied to nested design of higher order

The two-stage nested effects model is:

$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{k(ij)}$ (For a case of two factors)

$i = 1,2, \dots, a \; j = 1,2, \dots, b \; k = 1,2, \dots, n$

$\mu \; is \; the \; over \; all \; mean$

$\alpha_i \; is \; the \; i^{th} factor \; of \; effect \; A$

$\beta_{j(i)} is \; the \; j^{th} effect \; of \; factor \; B \; nested \; within \; the \; i^{th} level \; of \; factor \; A,$

$\epsilon_{ijk} is \; the \; random \; error \; of \; the \; k^{th} observation \; from \; the \; j^{th} level \; of \; B \; within \; the \; i^{th} \; level \; of \; A$

There are a levels of factor A, b levels of factor B nested under each level of A and n replicates. The subscript $j(i)$ indicates that the j^{th} level of factor B is nested under the i^{th} level of factor A. It is convenient to think of the replicates as being nested within the combination of levels of A and B; thus, the subscript $(ij)k$ is used for the error term.

2.1 Estimation of model parameters and Variance component

When the factors are fixed factors and following restriction are imposed on the estimates of the parameter.

$$\sum_{i=1}^{a} \hat{\alpha}_i = 0 \;\&\; \sum_{j=1}^{b} \hat{\beta}_{j(i)} = 0$$

Then the least square estimates of the model parameters are:

$\hat{\mu} = \hat{Y}_{...}$,

$\hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y}_{...}$ i = 1,2, . . . ,a and

$\hat{\beta}_{j(i)} = \bar{Y}_{ij.} - \bar{Y}_{i..}$ i = 1,2, . . . ,a and j = 1,2, . . . ,b

To estimate variance components, using the ANOVA method yields the following equation:

$\hat{\sigma} = MS_E \quad \hat{\sigma}_\alpha{}^2 = \frac{MS_A - MS_{B(A)}}{bn} \; \hat{\sigma}_\beta{}^2 = \frac{MS_{B(A)} - MS_E}{n}$

Random factors

Let $\boldsymbol{f}$ denote the effect size parameter

For a random factor at first level (say factor A)

$$\boldsymbol{f} = \sqrt{\frac{\hat{\sigma}_\alpha^2}{\sigma_A^2}}$$

For a random factor at second level (say factor B)

$$\boldsymbol{f} = \sqrt{\frac{\hat{\sigma}_\beta^2}{\sigma_B^2}}$$

Fixed factors

For a fixed factor at first level (say factor A)

$$\boldsymbol{f} = \sqrt{\frac{\sum_{i=1}^{a} \alpha_i^2}{\sigma_A^2\, bn}}$$

& For a fixed factor at second level (say factor B)

$$\boldsymbol{f} = \sqrt{\frac{\sum_{j=1}^{b} \beta_j^2}{\sigma_B^2\, n}}$$

Where,

$\hat{\sigma}_\alpha{}^2$ = variance component of factor A

$\hat{\sigma}_\beta{}^2$ = variance component of factor B

$\alpha_{i\,=}$ Least square estimate of model parameters for factor A

$\beta_{j\,=}$ Least square estimate of model parameters for factor B

$\boldsymbol{\sigma_A^2} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{i..} - y_{ijk})^2}{N-a}$ or

$\boldsymbol{\sigma_B^2} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{ij.} - y_{ijk})^2}{N-ab}$ (two-stage)

3.1 ILLUSTRATION

We will use a simulated data set to illustrate how to apply effect size for nested design for both random and fixed factors.

In this Illustration factor A is assumed to be fixed therefore, $\sum_i^a \alpha_i = 0$ for i = 1,2,...,a. That is, the A treatment effects sum to zero. While factor B is assumed to be random therefore, $\beta_{j(i)}$ is NID $(0, \sigma_\beta^2)$.

Factor A1				**Factor A2**			
Factor B1	Factor B2	Factor B3	Factor B4	Factor B1	Factor B2	Factor B3	Factor B4
82	79	78	79	85	86	83	84
83	78	82	83	85	84	83	85
79	81	83	80	83	85	85	85

Factor A3			
Factor B1	Factor B2	Factor B3	Factor B4
91	90	90	89
89	91	91	90
88	89	91	88

Factor A effect size

Factor A is a fixed factor therefore,

$f = \sqrt{\frac{\sum_{i=1}^{a} \alpha_i^2}{\sigma_A^2 bn}}$

$f = \sqrt{\frac{0.263928}{2.063131*4*3}}$

$f = 0.1305$

$f \approx 0.13$

Factor B is a random factor hence

$f = \sqrt{\frac{\hat{\sigma}_\beta^2}{\sigma_B^2}}$

$f = \sqrt{\frac{13.9784}{2.388889}}$

$f = 2.418972$

$f \approx 2.41$

Table 3.1: (Nested ANOVA Output from NeDPy)

ANOVA TABLE					
S.V	**D.F**	**SS**	**MS**	**F-RATIO**	**F-TAB**
Factor A	2.00	3.16670000	1.58330000	0.033100	4.256500
Factor B	9.00	430.916700	47.8796000	8.054500	2.300200
Error	24.0	142.666700	5.94440000		
Total	35.0	576.750000			

CONCLUSION

At alpha value of 0.05

Factor A is Insignificant

Factor B is Significant

3.1 DISCUSSION OF RESULT

Based on the extended formula, Factor A which is a fixed factor has an effect size of 0.13, while Factor B which is a random factor has an effect size of 2.41. Based on the proposed Cohen's benchmark, $\boldsymbol{f}$ = 0.10 -*small* effect, $\boldsymbol{f}$ = 0.25- *medium* effect, $\boldsymbol{f}$ = 0.40- *large effect*. Factor A is a small effect while Factor B is a medium effect. Inadvertently, these answers seemly corresponds with the Nested ANOVA output, as the Factor A is Insignificant at $\alpha = 0.05$, while Factor B is Significant at $\alpha = 0.05$.

4.0 CONCLUSION

These formulae are obviously easy and can effectively estimate the effect size in any nested data structure. They can be reported alone or reported alongside significance test to enrich the result thus help draw more accurate inferences. This computation can be carried out procedurally using most computational statistical software, although NeDPy II provides a menu that will directly compute effect size estimate using these procedures. At the point of this publication, the NeDPy version II has not been completed due to financial incapability. I solicit for support from all research lovers, those interested in the furtherance of this work and who find this work useful. These formulae can be extended to nested design of higher order.

Biographical Notes

Mr Marcel Obinna Tochukwu is a young 22 years old graduate. He earned his BSc (Mathematics & Statistics) from the University of Port-Harcourt (2016-2020). He has published in International journals and is currently carrying out several research awaiting Publication. He is the lead developer of the NeDPy statistical software. His Area of Interest is Experimental design and Statistical Computing. He aspires to pursue a higher degrees in Statistics.

Dr. Ogoeke earned B.Sc (Ed) (Mathematics) from the University of Nigeria, Nsukka, MSc and Ph.D degrees in statistics from the University of Port Harcourt. Presently she is a senior Lecturer in the Department of Mathematics and Statistics, University of Port Harcourt. She has attended International Workshops and Conferences where she presented her works and won a number of awards. She has published widely in both Local and International journals. She is also a member of relevant professional bodies such Nigerian Statistical Association, Nigeria Mathematical Society, International Biometric Society, Washington DC, USA and International Statistical Institute. She has her research interest in the area of Biostatistics.

REFERENCES

- Alf, E., & Abrahams, N. M. (1968). Relationship between percent overlap and measures of correlation. *Educational and Psychological Measurement*, *28*, 779-792.
- American Psychological Association. (2001). Publication manual of the American Psychological Association (5th ed.). Washington, DC: Author.
- Cohen, J. (1988). Statistical power analysis for the behavioural sciences (2nd ed.). New York: Academic Press.
- Dunnette, M. D. (1966). *Personnel selection and placement*. Belmont, CA: Wadsworth.
- Edgeworth, F. W. (1892). Correlated averages. *Philosophical Magazine* (5th series), *34*, 190-204.
- Galton, F. (1988). Co-relations and their measurement. *Proceedings of the Royal Society of London*, *45*, 135-145.
- Huberty, J. C. (2002). A History of Effect Size Indices Educational and Psychological Measurement, 62(2), 227-240.
- Johnson, H. D. (1999). The Insignificance of Statistical Significance Testing. *The Journal of Wildlife Management*, 63(3), 763-772.

- Konstantopoulos, S. (2008a). The Power of the Test in Three-Level Designs. Journal of Research on Educational Effectivness. Northwestern University.
- Marcel, T. O., Uchenna, P. O., & Ethelbert, C. N. (2020). Automation of Balanced Nested Design; NeDPy. *International Journal of Statistics and Applications*, 10(1), pp 17-23.
- Steiger, J. H., & Fouladi, R. T. (1997). Noncentrality interval estimation and the evaluation of statistical models. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), What if there were no significance tests? (pp. 221–257). Mahwah, NJ: Erlbaum.
- Sokal, R. R., & Rohlf, F. J. (1995). Biometry, 3rd ed. W. H. Freeman, New York. (1st ed. 1969; 2nd ed. 1981.)
- Tilton, J. W. (1937). The measurement of overlapping. *Journal of Educational Psychology*, *28*, 656-662.
- Wilkinson, L., & APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594–604.