

A comparative study between ridity and modified ridity analysis

Ebuh Godday Uwawunkonye, Oyeka Ikewelugo Cyprian Anaene

Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Nigeria

Email address:

ablegod007@yahoo.com (Ebuh G.U.)

To cite this article:

Ebuh Godday Uwawunkonye, Oyeka Ikewelugo Cyprian Anaene. A Comparative Study between Ridity and Modified Ridity Analysis. *American Journal of Theoretical and Applied Statistics*. Vol. 2, No. 6, 2013, pp. 248-254. doi: 10.11648/j.ajtas.20130206.23

Abstract: This paper compares ridity analysis with modified ridity analysis. The comparison was then illustrated with an example. It was observed from the example at least, that when the sample sizes of the two samples being compared are too disparate, a more reliable conclusion using the Bross ridity analysis is likely to be reached only when the group with the larger sample size is used as the reference group. Otherwise Bross ridity analysis would lead to conflicting conclusions, depending on which group is used as the reference group. Modified ridity analysis treats the groups being studied as samples drawn from some larger populations in which the variances or standard deviations as well as the results obtained are the same no matter which sample is used as the reference group. The modified procedure is therefore preferable to ridity analysis especially in cases where the groups being compared are samples from some populations.

Keywords: Samples Estimates, Populations, Bross Mean Ridity, Chi-Square, Significance

1. Introduction

Suppose that we have sample data drawn from a number of populations each of which is assumed to have in-built qualitatively ordered categories or classes. For example suppose we have random samples of patients by age say of a certain disease whose condition is ordered from critical to severe, poor, improved, most improved, etc. In an automobile accident involving some passengers, a passenger's level of injury may range from none through mild, severe to fatal.

Although these graduations may be coarse, discrete and still finite, they are never-the-less more descriptive and exhaustive than merely using some dichotomous classifications such as none or all, yes or no, present or absent, etc, which are fairly crude and not fully descriptive.

To compare these samples and reach clear conclusion is often difficult. However in the above and similar cases, the grading of the degree of seriousness is subjective and may not be reliable. Furthermore, it is difficult to find a readily interpretable summary index for such a data-set and to make comparisons among different samples in an intelligent way. The conventional chi-square analysis may be performed, but important information on the natural ordering of the categories would be lost.

A frequently employed procedure is to number the categories from say, 0 for the least serious to some highest

number for the most serious, calculate means and standard deviations, and then apply the 't' test or analysis of variance. There is however also a problem with this approach. The assignment of ordered numerical codes with equal spacing to the various categories of the variable under study is often arbitrary. It is a device that defines a metric on the categories of a qualitative variable which may or may not represent the true pattern of relationships among these categories.

A technique that does not attempt to quantify the categories but rather works with their natural ordering is the ridity analysis developed by Bross(1958).

The term ridity is an acronym for 'relative to an identified distribution' of the proportions or frequencies over the various ordered categories of some chosen standard or reference population, 'relative', in the sense that the proportions or frequencies of occurrence of observations in the various ordered categories of a population of interest, are compared with the proportions or frequencies in the corresponding ordered categories of the reference population or group. Virtually the only assumption made in ridity analysis is that the discrete categories represent intervals of an underlying but unobservable continuous distribution. No assumption is made about normality or any other form for the distribution.

In this paper, we briefly discuss ridity analysis, present a modified procedure, and use data to illustrate and compare the two techniques.

2. Ridit Analysis

Ridit analysis begins with the selection of one of the groups of data with the specified ordered categories to serve as a standard or reference population for the other groups, often referred to as comparison groups. Having selected a reference group or population, one then calculates a 'ridit' or score for each of its categories. The score or ridit for a given category is calculated as the cumulative frequency of all the categories lower in degree of seriousness than the category of interest plus one-half of the frequency for that category, all divided by the total frequency or the population size of the reference group. Thus using the data in the form of frequencies shown in table 1, the ridit for a category of the reference population is the proportion of all subjects or observations from the reference group falling in the lower ranking categories plus half the population falling in the given category

2.1. Methodology

Table 1: Data Format for Ridit Analysis

GROUPS

Ordered category of criterion variable (Ci)	Y (Reference, f_{iy})	X (Comparison, f_{ix})	Total (t_{ixy})
C ₁	f_{1y}	f_{1x}	$t_{1xy}(=f_{1x}+f_{1y})$
C ₂	f_{2y}	f_{2x}	$t_{2xy}(=f_{2x}+f_{2y})$
.	.	.	.
.	.	.	.
.	.	.	.
C _k	f_{ky}	f_{kx}	$t_{kxy}(=f_{kx}+f_{ky})$
Total	n_y	n_x	n_x+n_y

Once the ridits for all the categories of the reference population are determined, they are taken as values of a dependent variable for the other groups. Given the distribution of any other group over the same categories the mean ridit for that group may be calculated. It is simply the sum of the products of observed frequencies, times the ridits obtained from the reference population for the corresponding categories divided by the total frequency for that group.

Thus using the frequencies in Table 1, the ridit of the i^{th} category of the reference population Y is

$$r_{iy} = \frac{\left(\sum_{j=1}^{i-1} f_{jy} + f_{iy} \right)}{n_y} = \sum_{j=1}^{i-1} p_{jy} + p_{iy} \quad (1)$$

Where p_{jy} and p_{iy} are respectively the proportions of the total observations in the j^{th} and i^{th} categories of the reference population Y for $j = 1, 2, \dots, i-1, i = 1, 2, \dots, k$. The mean ridits \bar{r}_x for any other group X is then calculated as

$$\bar{r}_x = \frac{\left(\sum_{i=1}^k f_{ix} r_{iy} \right)}{n_x} = \sum_{i=1}^k p_{ix} r_{iy} \quad (2)$$

where p_{ix} is the proportion of cases in, or the relative frequency of the i^{th} category of group or population X, for $i = 1, 2, \dots, k$. The mean ridit for the reference population Y is by the definitions in equations 1-2 always equal to 0.50. This means that if any two subjects are selected at random from the reference population Y, then one of them would be expected to experience a more serious condition on the criterion variable half of the time, and a less serious condition also half of the time than the others subject in the reference population.

The mean ridit for any other group X, is interpreted as follows; Given the reference group Y and any other group X, then the mean ridit \bar{r}_x for the comparison group X is an estimate of $P(X \geq Y)$, that is of the probability that a randomly selected subject from group X, the comparison group, has a condition that is at least as serious as that of a randomly selected subject from group Y the reference group on the criterion variable. Thus if the mean ridit \bar{r}_x for a given comparison group X is more than 0.50, then more than half of the time or randomly selected subject from it will have a more serious condition than a randomly selected subject from the reference group Y. If on the other hand the ridit for the group is less than 0.50, we would conclude that a randomly selected subject from it would be expected to experience a less serious condition than a randomly selected subject from the reference group Y. A mean ridit of 0.50 for any group would imply that subjects from that group would tend to experience neither more nor less serious condition than subjects from the reference group.

Therefore if R_x is the mean ridit of a comparison population X from which a random sample of size n_x has been drawn to obtain \bar{r}_x , then a null hypothesis that needs to be tested is

$$H_0: R_x \geq 0.50, \text{ versus } H_1: R_x \leq 0.50 \quad (3)$$

Where 0.50 is the mean ridit of the standard or reference population Y.

It has been shown by Bross(1958) that for sufficiently large sample size n_x , \bar{r}_x is approximately normally distributed with mean R_x and variance

$$Var(\bar{r}_x) = \frac{1}{12n_x} \quad (4)$$

Hence the null hypothesis of Eqn 3 may be tested using the test statistic.

$$\chi^2 = \frac{(\bar{r}_x - 0.50)^2}{Var(\bar{r}_x)} = 12n_x (\bar{r}_x - 0.50)^2 \quad (5)$$

Which has approximately a chi-square distribution with 1 degree of freedom for sufficiently large n_x . H_0 is rejected at the α level of significance if

$$\chi^2 \geq \chi_{1-\alpha;1}^2 \quad (6)$$

Otherwise H_0 is accepted.

2.2. Illustrative Example 1

The data of Table 2 shows the degree of the effects of the concentration of some poisonous chemical on the blood stream of three groups of employees by work place.

Table 2: Distribution of Employees by work place and level of reaction of some poisonous chemical

Work place

Reaction to poisoning (severity of condition)	A(Y)	B(X)	C(Z)	Total		
				T _{xy}	T _{yz}	T _{xz}
None	35	58	24	93	59	82
Moderate	31	392	126	423	157	182
Severe	17	96	39	113	56	135
Critical	8	21	29	29	37	50
Fatal	14	48	24	62	38	72
Total	105(=n_y)	615(=n_x)	242(n_z)	720	347	523

It can be seen from Table 2 that groups Y and X have a sample size ratio of approximately 1:6, groups Y and Z, a sample size ratio of approximately 1:2 and groups X and Z a sample size ratio of approximately 3:1. To find out the effect different sample sizes have on ridit analysis, we have use each of the groups or populations X, Y and Z alternatively as the comparison and reference populations.

First using Y(A) as the reference group we calculate the ridits for the categories of Y from Table 2 using Equation 1 as

$$\text{None: } r_{1y} = \frac{35/20}{105} = 0.167;$$

$$\text{Moderate: } r_{2y} = \frac{35+31/2}{105} = 0.481;$$

$$\text{Severe: } r_{3y} = \frac{35+31+17/2}{105} = 0.710;$$

$$\text{Critical: } r_{4y} = \frac{35+31+17+8/2}{105} = 0.829;$$

$$\text{Fatal: } r_{5y} = \frac{35+31+17+8+14/2}{105} = 0.933;$$

From the ridit scores we calculate the mean ridit for X(B) Equation 2 as

$$\bar{r}_x = \frac{(58)(0.167) + (392)(0.481) + (96)(0.710) + (21)(0.829) + (48)(0.933)}{615}$$

$$\bar{r}_x = \frac{328.5316}{615} = 0.534$$

Thus the ridit analysis estimates that the probability is 0.534 that a randomly selected employee in work place B(X) has as serious or more serious reaction to the chemical poisoning than a randomly selected employee in work place A(Y), the reference group.

The corresponding variance is from Equation 4

$$Var(\bar{r}_x) = \frac{1}{12(615)} = \frac{1}{7380} = 0.0001$$

Hence the test statistic for the null hypothesis of Equation 3 is

$$\chi^2 = \frac{(0.534 - 0.50)^2}{0.0001} = 11.56$$

which with 1 degree of freedom is highly statistically significant.

A similar calculation may be made for work place C(Z) as a comparison group again using work place A(Y) as the reference group yielding an estimated mean ridit.

From Equation 2 as

$$\bar{r}_z = \frac{138.7059}{242} = 0.573$$

with a variance of

$$Var(\bar{r}_z) = \frac{1}{12(242)} = 0.0003$$

Hence the corresponding chi-square test statistics is

$$\chi^2 = \frac{(0.573 - 0.50)^2}{0.0003} = \frac{0.0053}{0.0003} = 17.763$$

which is again highly statistically significant.

If instead work place B(X) has been used as the reference group or population the ridit scores for X would be

$$\text{Mild: } r_{1x} = \frac{58/2}{615} = 0.047;$$

$$\text{Moderate: } r_{2x} = \frac{58+392/2}{615} = 0.413;$$

$$\text{Severe: } r_{3x} = \frac{58+392+96/2}{615} = 0.810;$$

$$\text{Critical: } r_{4x} = \frac{58+392+96+21/2}{615} = 0.905;$$

$$\text{Fatal: } r_{5x} = \frac{58+392+96+21+48/2}{615} = 0.961$$

Using these values in Equation 2 we calculate the mean ridit for work place A(Y), now treated as a comparison

$$\text{group as } \bar{r}_y = \frac{48.915}{105} = 0.466$$

This value is simply equal to $1 - \bar{r}_x$ which is meaningfully reflecting the fact that a complementary probability is being estimated. The variance of this mean ridit from Equation 4 is

$$Var(\bar{r}_y) = \frac{1}{12(105)} = 0.0008$$

The corresponding chi-square test statistics is

$$\chi^2 = \frac{(0.466 - 0.50)^2}{0.0008} = 1.445$$

which is not significant at the 5% significance level. Thus the change in choice of the reference group or population has resulted in a non significant effect. In other words, if we had used work place B(X) as the reference group or population instead of work place A(Y) we would conclude that employees in work place A(Y) now treated as the comparison group are as seriously affected by the chemical poisoning as the employees in work place B(X). But if we had used workplace A(Y) as the reference group we would conclude that employees in workplace B(X) are more seriously affected by the chemical poisoning than the employees in workplace A(Y).

Hence conclusions reached using ridit analysis often depend on which group is used as a reference group or population, and which is used as a comparison group.

We now present a modified method of estimating ridits that are independent of which populations are used as reference and which as comparison groups.

3. Modified Ridit Analysis

Implicit in the Bross ridit procedure is the assumption that the reference group is a population. Although the author did mention the difficulty of selecting an appropriate reference group, he failed to explicitly suggest an appropriate procedure when either of the two groups to be compared might serve as a reference group. The main cause of the problem with the Bross procedure is in the difficulty of determining an appropriate standard deviation to use in the denominator of the test statistic. Interchanging the reference and comparison groups merely interchanges the roles of these groups, while the mean ridits estimated are still meaningful and useful probabilities. However, if the sizes of the two groups that are being compared are very different using one of the groups rather than the other as a reference group affects the standard deviation and hence result of the test. Furthermore, if all available groups are regarded as samples from their respective populations, an additional source of variability is also introduced, since the ridit scores are then subject to variations themselves.

The results of the following procedure are similar to those of ridit analysis in interpretation but the procedure makes the explicit assumption that all the groups are to be regarded as samples from their respective populations. The method is based on Mann and Whitney (1947). Area works by Conover (1973) and Oyeka (1992) provide theoretical bases. Other researches include Mieke et al (2009), Pouplard et al (1997), and Rao and Caliguin (1993).

3.1. Methodology

Let $y_{j(j=1,2,\dots,n_y)}$ be a set of n_y observations made on group or population Y (Bross reference group), occurring with frequencies, $f_{1y}, f_{2y}, \dots, f_{ky}$ across c_1, c_2, \dots, c_k the k categories of a criterion variable such that $n_y = \sum_{i=1}^k f_{iy}$.

Similarly let $x_{i(i=1,2,\dots,n_x)}$ be a set of n_x observations made on any other group or population X referred to as the comparison group occurring with frequencies $f_{1x}, f_{2x}, \dots, f_{kx}$ respectively across the k categories of the same criterion variable, such that $n_x = \sum_{i=1}^k f_{ix}$.

The data format is as in Table 1

Now for the reference group Y and any comparison group X define the function u_{ij} as

$$u_{ij} = \begin{cases} 1, & \text{if } x_i > y_j \\ 0, & \text{if } x_i = y_j \\ -1, & \text{if } x_i < y_j \end{cases} \quad (7)$$

for $i = 1, 2, \dots, n_x; j = 1, 2, \dots, n_y$

Let

$${}^+ \pi = P(u_{ij} = 1): {}^0 \pi x = P(u_{ij} = 0): {}^- \pi x = P(u_{ij} = -1) \quad (8)$$

$$\text{Where } {}^+ \pi x + {}^0 \pi x + {}^- \pi x = 1 \quad (9)$$

Also define

$$W_x = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} u_{ij} \quad (10)$$

Now

$$E(u_{ij}) = {}^+ \pi x - {}^- \pi x; Var(u_{ij}) = {}^+ \pi x + {}^- \pi x - \left({}^+ \pi x - {}^- \pi x \right)^2 \quad (11)$$

Also

$$E(W_x) = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} E(u_{ij})$$

OR

$$E(W_x) = n_x n_y \left({}^+ \pi x - {}^- \pi x \right) \quad (12)$$

Note that ${}^+ \pi x$, ${}^0 \pi x$ and ${}^- \pi x$ are respectively the

probabilities that a randomly selected subject from the comparison population X is in a more serious, as serious as or less serious condition on the criterion variable than a randomly selected subject from the reference population Y. In terms of Bross mean ridit

$$\bar{r}_x = \pi_x^+ + \frac{1}{2} \pi_x^0; 1 - \bar{r}_x = \pi_x^- + \frac{1}{2} \pi_x^0 \quad (13)$$

These probabilities are estimated as functions of f_x^+ , f_x^0 and f_x^- where f_x^+ , f_x^0 and f_x^- are respectively the number, or frequencies of occurrence, of 1's 0's and -1's in the frequency distribution of the $n_x n_y$ values of these numbers in u_{ij} , $i = 1, 2 \dots n_x$; $j = 1, 2 \dots n_y$.

The sample estimates of $\pi_x^+ - \pi_x^-$, the difference between π_x^+ and π_x^- is from equation 12

$$\pi_x^+ - \pi_x^- = \frac{W_x}{n_x n_y} = \frac{f_x^+ - f_x^-}{n_x n_y} \quad (14)$$

Also the sample estimate of the variance of w_x is obtained using equation 13 as

$$Var(w_x) = n_x n_y \left(\hat{\pi}_x^+ + \hat{\pi}_x^- \right) - \frac{W_x^2}{n_x n_y} \quad (15)$$

Hence the sample estimate of π_x^0 is

$$\hat{\pi}_x^0 = \frac{f_x^0}{n_x n_y} \quad (16)$$

Also it is easy to show using equations 9 and 14 that the estimated values of π_x^+ and π_x^- are respectively

$$\hat{\pi}_x^+ = \frac{1}{2} \left(1 - \hat{\pi}_x^0 + \frac{W_x}{n_x n_y} \right) = \frac{1}{2} \left(1 - \frac{f_x^0}{n_x n_y} + \frac{(f_x^+ - f_x^-)}{n_x n_y} \right) \quad (17)$$

And

$$\hat{\pi}_x^- = \frac{1}{2} \left(1 - \hat{\pi}_x^0 - \frac{W_x}{n_x n_y} \right) = \frac{1}{2} \left(1 - \frac{f_x^0}{n_x n_y} - \frac{(f_x^+ - f_x^-)}{n_x n_y} \right) \quad (18)$$

It has been shown by Hajek (1969) and Conover (1973) that the variance of

W_x is estimated as

$$Var(W_x) = \frac{n_x n_y (n_x + n_y + 1)}{3} \left[1 - \frac{\sum_{i=1}^k (t_{ixy}^3 - t_{ixy})}{(n_x + n_y)^3 - (n_x + n_y)} \right] \quad (19)$$

that a randomly selected subject from the comparison population X_i is in a more serious condition on the criterion variable than a randomly selected subject from the reference population Y and the probability that the randomly selected subject is in a less serious condition, provides a measure of the relative seriousness of the condition in the populations. Now if population X and population Y are in fact the same population, then the relative difference between the probabilities of seriousness of the condition in the two populations would be = Zero, so that $\pi_x^+ - \pi_x^- = 0$

Hence a more general null hypothesis that needs to be tested here is

$$H_0: \pi_x^+ - \pi_x^- \geq \delta_0 \text{ versus } H_1:$$

$$\pi_x^+ - \pi_x^- < \delta_0, (-1 < \delta_0 < 1) \quad (20)$$

Now for sufficiently large n_x and n_y , W_x has approximately the normal distribution with mean $E(W_x)$ of equation 12 and variance, $Var(W_x)$ of equation 19.

Hence the null hypothesis of equation 20 may be tested using the test statistic

$$\chi^2 = \frac{(W_x - E(W_x))^2}{Var(W_x)} \quad (21)$$

OR

$$\chi^2 = \frac{(W_x - n_x n_y \delta_0)^2}{Var(W_x)} \quad (22)$$

Which has approximately a chi-square distribution with 1 degree of freedom for sufficiently large n_x , n_y where $Var(W_x)$ is given in equation 19. H_0 is rejected at the α level of significance if Equation 6 is satisfied. Otherwise H_0 is accepted.

3.2. Illustrative Example 2

Let us now use the modified method to re-analyze the data in Table 2 earlier used to illustrate ridit analysis in example 2. We first use work place A(Y) as the reference population and work place B(X) as the comparison population

Now from equation 13 and Table 2, we have that

$$f_x^+ = (392)(35) + (96)(35+31) + (21)(35+31+17) + (48)(35+31+17+8) \\ = 26167$$

Also

$$f_x^- = (58)(31+17+8+14) + (392)(17+8+14) + (26)(8+14) + (21)(14) \\ = 21754$$

$$\text{And } f_x^0 = (58)(35) + (392)(31) + (96)(17) + (21)(8) + (48)(14) \\ = 25954$$

Hence from equation 14 we have that

$$\hat{\pi}_x - \pi_x = \frac{4413}{64575} = 0.068$$

$$\text{Also from equation 3.13 } \hat{\pi}_x = \frac{25954}{(615)(105)} = 0.402$$

From equation 17

$$\hat{\pi}_x = \frac{1}{2}(1 - 0.402 + 0.068) = 0.333$$

And from equation 18

$$\hat{\pi}_x = \frac{1}{2}(1 - 0.402 + 0.068) = 0.265$$

Therefore the estimated probabilities are

$$\hat{\pi}_x = 0.333; \hat{\pi}_x = 0.402 \text{ and } \hat{\pi}_x = 0.265$$

Hence the modified approach estimates that the probability is 0.333 that a randomly selected employee in work place B(X) (the comparison population) is more seriously affected by the chemical poisoning than a randomly selected employee in work place A(Y) (the reference population) 0.402 that the employee is as seriously affected and 0.265 that the employee is less seriously affected.

Notice that as earlier obtained using Gross mean ridit

$$\bar{r}_x = 0.333 + \frac{1}{2}(0.402) = 0.534$$

The fact that the probability of experiencing equal severity of condition has now been estimated separately is an important and useful additional information and advantage of the modified method over Bross ridit analysis.

The variance of W_x is estimated from equation 19 as

$$\text{Var}(W_x) = \frac{(615)(105)(721)}{3} \left[1 - \frac{78196218}{373247000} \right] = 12,268,144$$

The null hypothesis of Equation 20 (with $\delta_0 = 0$) may now be tested using the test statistic

$$\chi^2 = \frac{(4413 - 0)^2}{12,268,144} = \frac{19,474,569}{12,268,144} = 1.587$$

which with 1 degree of freedom is not statistically significant. If we now interchange the roles of Y and X, that is if X now becomes the reference group and Y, becomes the comparison group, we would have that

$$\begin{aligned} f_y^+ &= \sum_{j=2}^k f_j y \left(\sum_{i=1}^{i-1} f_i x \right) \\ &= 31(58) + 17(58+392) + 8(58+392+96) + 14(58+392+96+21) \\ &= 26167 \end{aligned}$$

And

$$\begin{aligned} f_y^- &= \sum_{j=2}^{k-1} f_j y \left(\sum_{i=j+1}^k f_i x \right) = \\ &= 35(392+96+21+48) + 31(96+21+48) + 17(21+48) + 8(48) \\ &= 21754 \end{aligned}$$

$$\text{Also } f_y^0 = \sum_{i=1}^k f_i y f_i x = 25954 \text{ as before}$$

Therefore,

$$\begin{aligned} \pi_y - \pi_y &= \frac{f_y^+ - f_y^-}{n_x n_y} = \frac{W_x}{n_x n_y} = \frac{21754 - 26167}{(615)(105)} = \frac{-4413}{(615)(105)} \\ &= -0.068 \end{aligned}$$

$$\text{And } \pi_y = \frac{25954}{(615)(105)} = 0.402$$

$$\text{Also } \pi_y = \frac{1}{2}(1 - 0.402 - 0.068) = 0.265$$

$$\text{And } \pi_y = \frac{1}{2}(1 - 0.402 + 0.068) = 0.333$$

$$\text{Hence we have that } \pi_y = 0.265; \pi_y = 0.402 \text{ and}$$

$$\pi_y = 0.333$$

These are the same values obtained when work place A(Y) is used as the reference population and work place B(X) is used as the comparison population. Hence the value of the estimated variance of W_y remains the same as that of W_x and the value of the test statistic and the attained significance level remain unchanged. Hence the same conclusions are reached with the modified ridit analysis irrespective of which of the populations is used as the reference group and which as the comparison group.

Thus the only changes that result when the roles of the two groups are interchanged are that the sign of W changes;

π_x is interchanged with π_y and π_x is interchanged

with π_y , unlike is the case with Bross ridit analysis.

Table 3 shows the results of the analysis of the data in Table 2 by both the ridit analysis and the modified ridit analysis using each of the work places in turn as the reference group or population

4. Results and Discussions

The chi-square test statistics of Equation 5 and 22 may be expressed in terms of unit-normal z-scores and are hence so reported in Table 3 as can be seen from this table, interchanging the reference and comparison groups has a marked effect on the Bross analysis, often changing the attained significance level drastically. In modified ridit analysis no such effect is observed. Here the variance (standard deviations) and the attained significance levels remain the same when the reference and comparison groups are interchanged. Furthermore the significance levels attained by the Bross ridit analysis appear to be generally lower than those by the modified ridit analysis. In fact it is only when the same size of the reference group is very large, as in group X, that the significance levels attained by the two procedures are approximately the same.

Thus, it would seem, from the above example at least, that when the sample sizes of the two samples being compared are two disparate, a more reliable conclusion using the Bross ridit analysis is likely to be reached only when the group with the larger sample size is used as the reference group. Otherwise Bross ridit analysis would lead to conflicting conclusions, depending on which group is

used as the reference group. With the modified ridit analysis on the other hand, no such problem is encountered. Furthermore when the difference between the sample sizes of the groups being compared, is not too large, as in the

case of groups X and Z, the attained levels of significance seem to suggest that the modified procedure would still be the preferred method in terms of consistency of conclusions.

Table 3: Comparison of two Ridits methods

Data	Ridit Analysis				Modified Ridit Analysis					
	\bar{F}	S	Z	p-value	π	π	π	S	Z	P-value
Workplace A(Y) as reference group										
Workplace B(X)	0.534	0.012	2.83	0.005	0.333	0.402	0.265	3502.60	1.26	0.208
Workplace C(Z)	0.573	0.019	3.84	0.00	0.456	0.235	0.309	1624.89	2.29	0.022
Workplace B(X) as reference group										
Workplace A(Y)	0.466	0.028	-1.21	0.226	0.265	0.402	0.333	3502.60	-1.26	0.208
Workplace C(Z)	0.554	0.019	2.84	0.005	0.365	0.378	0.257	5738.39	2.80	0.005
Workplace C(Z) as reference group										
Workplace A(Y)	0.427	0.028	-2.61	0.009	0.309	0.235	0.456	1624.89	-2.29	0.022
Workplace B(X)	0.448	0.012	-4.33	0.000	0.257	0.378	0.365	5738.35	-2.80	0.005

5. Conclusion

Modified ridit analysis treats the groups being studied as samples drawn from some larger populations in which the variances or standard deviations and the results obtained are the same no matter which sample is used as the reference group. The modified procedure is therefore preferable to ridit analysis especially in cases where the groups being compared are samples from some populations.

References

- [1] Bross, DJ. (1958) How to Use Ridit Analysis. Biometrics 14 Pgs18-38.
- [2] Mann, HB. and Whitney, DR. (1947) On a Test of whether one of two random variables is stochastically larger than the other. The Annals of Mathematical Statistics 18(1):50-60
- [3] Conover, WJ. (1973) Rank Tests for One Sample, Two Samples and K Samples Without the Assumptions of a Continuous Distribution Function. Annals of Statistics 1. Pgs 1105-1125
- [4] Oyeka, ICA. (1992) A Modified Ridit Analysis. Journal of Nigerian Statistical Association (JNSA) Vol8, No 1 Pgs 45-69
- [5] Mieke, PW. Jnr; Long, MA.; Berry, KJ.; and Johnson, JE.(2009) g-Treatment ridit analyses:Resampling Permutation methods Statistical Methodology, Volume 6, Issue3, May 2009, Pgs 223-229.
- [6] Pouplard, N; Quannari, EM.; and Simon, SC.(1997)Use of ridits to analyse Categorical data in preference studies .Food Quality and Preference Volume 8,Issues 5-6,September-November 1997, Pgs 419-422 Third Sensometrics Meeting. Doi: 10.1016/S0950-3293(97)00020-7
- [7] Rao, CR. and Caliguin, MP. (1993) Analysis of ordered Categorical Data through appropriate Scaling. Handbook of Statistics Volume 9, 1993, Pgs 521-533. Computational Statistics. Doi: 10.1016/S0169-7161(05)80139-7
- [8] Hajek, J. (1969) Nonparametric Statistics. Holden-Day, San Francisco.