SEVERITY ANALYSIS USING RIDITS

Modified in 2023

Abstract

The technique of ridit analysis used in severity analysis was studied for its feasibility for use in environmental toxicity studies. The need for an alternative to the t-test and chi-square families of techniques arises because an ordered series of categories frequently cannot be laid out on an equal interval scale (Kantor, 1967, page 610).

1. Introduction

A generic SAS program was developed for use with multiple dose levels to calculate the ridits, mean ridits by group, the overall population mean ridit and a chi-square statistic to test the hypotheses of no dose difference with respect to the control dose. The reference group was the population although a control group can also be specified.

2. Statistical Theory

Ridit analysis is proposed as the method of choice for analyzing ordered categorical data. Bross [1958] developed the use of Ridit analysis for ordinally scaled data such as injury severity categories. A ridit is a numerical quantity (0<ridit<1); it is a transformation converting a severity classification into a number.

The only assumption made in ridit 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. Ridit analysis begins with the selection of a population to serve as a standard or reference group.

We know that a given percentile is that value which divides the range of a set of data into two parts such that a given percentage of the measures lies below this value. It is therefore a probability. Ridit analysis transforms ordinal data to a probability scale. The ridit score for each category is a percentile rank of a subject in the reference population and is equal to the number of subjects in all lower categories plus one-half the number of items in the subject category, all divided by the population size. The score (ridit) given to a severity category is the relative frequency up to the midpoint of that category in the reference group.

Once the ridits for each category have been determined, they are considered as a dependent variable for the other comparison groups (Jairus Flora, 1974, page 2) and the usual normal probability distribution family of statistics can be applied in calculating means, standard deviations etc. The mean ridits calculated in this way will be approximately normal for reasonable sample sizes.

The mean ridit for the comparison group is determined as follows. If a subject X, is selected at random from the reference population (control

group) and a subject Y, is selected at random from the comparison group, then the mean ridit is an estimate of P(X≤Y), that is, of the probability that X is less seriously injured than Y. The control group mean ridit is always .5 under this definition.

Let the reference group be group 1. This is in conformance to the SAS program RIDITS referenced in the bibliography. Let P_{ij} be the proportion in severity category j=1...,k of the group I and define the ridit for a severity category by

$$R_j = \sum_{n=1}^{j-1} P_{1n} + P_{1j}/2 \qquad (2.1)$$

If X denotes the injury severity for a subject selected at random from the reference population, and Y denotes the injury severity for a subject within a particular group (dose), then the mean ridit for that group (dose)

$$\bar{R}_i = \sum_{j=1}^{k} R_j \, P_{ij} \qquad (2.2)$$

can be interpreted as an estimate that a subject from the reference group would be in a less severe severity classification code than a dosed subject. More precisely,

$$\bar{R}_i \text{ estimates } P[X<Y] + \tfrac{1}{2} P[X=Y] \qquad (2.3)$$

The population mean ridit R is calculated as

$$R = \frac{\sum_{1}^{k} R_i N_i}{\sum_{1}^{k} N_i} \qquad (2.4)$$

It is the probability that for a subject selected at random from the population the subject would be in a less severe severity code level than a subject which has been dosed by a pesticide.

The standard error (s.e. or SE) of the mean ridit of a group of size N_i is approximately (Fleiss, 1987, p.105)

$$\frac{1}{2\sqrt{3N_i}} \qquad (2.5)$$

The usual confidence interval for a mean ridit (Bross, 1956, page 7) is calculated as

$$R_i \pm z_{\alpha/2} \sqrt{\frac{R_i \ (1 - R_i)}{(N_i - 1)}} \qquad (2.6)$$

The normal pdf statistic may be replaced by a t statistic in the case of small sample sizes when constructing confidence intervals. Rough 95% confidence intervals on ridit means (Bross, 1958, page 24) are calculated as

$$R_i \pm 1/\sqrt{3 \ N_i} \qquad (2.7)$$

Confidence intervals on average ridits involve the probability of a probability statement being true; it is a probability on the odds (Bross, 1956, page 8). The odds statement is expressed by the ratio (Kantor, 1968, page 613)

$$\frac{0.50 + d}{1 - (0.50 + d)} \qquad (2.8)$$

where d is the numerical difference between the average ridits of the groups being compared. Ridit differences are declared statistically significant when the 95% confidence limits of the two groups being compared do not intersect.

When comparing two groups without involving the reference group add .5 to the numerical differences of the two mean ridits (Bross, 1956, page 8). Thus if R_i R_j are the mean ridits for groups I and j, $R_i - R_j + 0.5$ estimates the probability that a randomly selected member of group i has a value on the underlying variable greater than or equal to that of a randomly selected member of group j.

For example, let $R_i - R_1 + 0.5 = .66$. In terms of ridits the odds are about 2 to 1 that a dosed individual will sustain worse injury than an individual randomly selected from the control group one.

The rectangular distribution on the interval from 0 to 1 has a mean of .5 and an estimated variance of 1/12. The variance for a difference in two mean ridits is given by:

$$V(R_i-R_j) = 1/12N_i+1/12N_j = \frac{N_j+N_i}{12N_iN_j} \qquad (2.9)$$

Between group comparisons on the basis of mean ridits use t-tests and z-tests. The standard error for the difference of two mean ridits is given by

$$s.e.(R_i-R_j)=\frac{\sqrt{()N_i+N_j}}{2\sqrt{(3N_i\ N_j)}} \qquad (2.10)$$

The significance of the difference between R_2-R_1 (individual contrast) for the large sample case may be tested by forming the statistic

$$z_{ij} = \frac{(R_i-R_j)}{s.e.(R_i-R_j)} = \frac{(R_i-R_j)}{\sqrt{N_i+N_j}/(2\sqrt{3N_iN_j})} = \frac{(R_i-R_j)2\sqrt{3N_iN_j}}{\sqrt{N_i+N_j}} \qquad (2.11)$$

The process of analyzing a sample data set using ridits is also analogous to ANOVA (Fleiss, 1986, page 81). An overall F-test of equal means is replaced by a chi-square test of "g equal mean ridits" with degrees of freedom varying as there are groups being compared. The test statistic is

$$\chi^2_{g-1,\alpha} = \frac{12N_i(\bar{R}_i-R_.)^2}{(N_.+1)f} \qquad (2.12)$$

where f is defined as

$$f=1-\frac{\sum_1^T t_i(t_i-1)(t_i+1)}{N.(N.-1)(N_i+1)} \qquad (2.13)$$

In the above equation, t_i represents the number of ties at the ith response. There are a total of T responses.

The significance of the difference between groups i and j may be tested simultaneously using a Scheffe`-type criterion. Compare the value of calculated X^2 with $X^2_{g-1,\alpha}$ where f_{ij} is the value of an adjustment factor calculated only on the frequencies of i and j combined. The test statistic is

GROUP COMPARISONS - Scheffe`

$$X^2_{g-1,\alpha} = \frac{12N_iN_j(R_i-R_j)^2}{(N_i+N_j+1)f_{ij}}$$

where f_{ij} *is defined as* (2.14)

$$f_{ij} = 1-\frac{\sum_1^g t_i(t_i-1)(t_i+1)}{n_.(n_.-1)(n_.+1)}$$

t_i *represents the number of ties at the ith response*

A typical Scheffe confidence interval testing the hypothesis of equal ridit means for groups i and j is given by

$$(R_i-R_j) \pm S\frac{\sqrt{N_i+N_j}}{2\sqrt{3N_iN_j}}, \text{ where } S = \chi^2_{g-1,\alpha/2} \qquad (2.15)$$

The probability is 1-α that the values of all contrasts simultaneously satisfy the inequalities $(R_i-R_j) + S*SE \leq (R_i-R_j) \leq (R_i-R_j) + S*SE$ (Scheffe, 1959, page 67.

A second approach to confidence interval construction utilizes the so called Bonferroni Inequality which has the following theoretical basis (Mendenhall, 1990, page 624). For any events $A_1, A_2, \ldots A_m$ we have

$$P(A_1 \cap A_2 \cap A_m) = 1 - P(P(A_1 \cup A_2 \cup \ldots A_m)\ldots \qquad (2.16)$$

From the additive law of probability we know that

$$P(A_1 \cup A_2 \cup \ldots A_m) \leq P(A_1) + P(A_2) + \cdots P(A_m) \qquad (2.17)$$

Therefore, it follows that

$$P(A_1 \cap A_2 \cap \ldots A_m) \geq 1 - (P(A_1) + P(A_2) + \cdots P(A_m)) \qquad (2.18)$$

Suppose that we want to find confidence intervals $I_1, I_2, \ldots, I_m$ each covering 0 for parameters $\theta_1, \theta_2, \ldots, \theta_m$ representing the m differences of treatment means minus control means where m = number of contrasts. Suppose that $P(\theta_j \in I_j) = 1-\alpha_j, \ \forall j$ and let A_j denote the event that θ_j is in I_j. Then

$$P(\theta_1 \in I_1, \ldots, \theta_m \in I_m) \geq 1 - \sum_1^m P(\theta_j \notin I_j)$$

$$= 1 - \sum_1^m \alpha_j = 1 - m\,\alpha \qquad (2.19)$$

if all α_j, $j=1,m$ *are chosen equal* α

A simultaneous confidence coefficient of at least $(1-\alpha)$ can be assured by choosing each $I_j, j=1, \ldots m$, to have confidence coefficient $1-(\alpha/m)$. This is a two-tailed test.

Adjust the level of significance (_) of each test to maintain a desired overall level α for the whole set of comparisons comprising the original null hypothesis. This goal can be achieved by using the so-called Bonferroni inequality with the α for an individual comparison is α/m where m is the number of comparisons.

A typical Bonferroni confidence interval (Fleiss, 1979, page 2082) testing the hypothesis of equal mean ridits for groups I and j is

$$(R_i - R_j) \ \pm \ B_{\alpha/2m} \frac{\sqrt{N_i + N_j}}{2\sqrt{3 N_i N_j}} \qquad (2.20)$$

where B is the corresponding critical normal curve value in the case of large sample sizes or the critical t curve value in the case of small sample sizes.

The mean ridit calculated for the reference population will always be **0.5** by definition.

Implicit is the assumption that the reference group is a population. Ranges on the variance of the ridit scores in the reference population (Flora, 1974, page 2) depend on the shape of the distribution being at most 1/4, about 1/12 for approximately equal numbers in the categories, and < 1/12 for extremely skewed distributions.

Variance of the ridit scores in the Comparison Group or variance(R_i) =1/4g, g=# of groups.

An ultra conservative test is

$$H_o: \ R_i = R_{Ref}, \quad t = \frac{(R_i - .5)}{\sqrt{1/4g}} \qquad (2.21)$$

3. Suggested Protocol

A suggested protocol for the execution of a ridit analysis for a dose/response data set is as follows.

Step 1

Rank the entire data sets response variable over all doses.

Step 2

Determine the severity code ranges based upon the ranks. For example consider the situation where there are 70 observations, 70/5=14. Now assign each observation a severity code using cutoff points 14,28,42,56,70. A sample table might look like

Severity Code Ranges (N=14)

Dose	1-14	15-28	29-42	43-56	57-70
0	6	7	3	2	2
3	2	2	4	5	5
6	5	3	4	2	5
12	1	2	3	5	2

Step 3

Perform a ridits analysis.

4. Analysis of Severity.

Two data sets were examined. The first was used to evaluate severity of glomerulonephrophy in male rat kidneys with dose increments of the pesticide X and a second evaluation of severity of mononuclear cell leukemia with dose increments of the pesticide Y.

On the basis of the severity distributions seen in Tables 1 and 2 below, a statistical analysis was calculated for each dose group in terms of its weighted ordered severity score. This technique developed by Bross in 1958 was called ridits (relative to an identified distribution). The score itself is called the mean ridit for a dose group.

Ridit analysis is a way of making comparisons among different samples of an endpoint selected by intervals. The endpoint selected in the studies in this paper is the number of individuals falling into specified severity codes (no disease, minimum, mild, moderate, and severe) by dose group samples. The intervals themselves consist of the severity codes.

The primary sampling unit is an individual; and from it is determined a level of severity based upon a continuous random variable which is divided into ordinal categories. Associated with each individual from the identified distribution is a numerical quantity (ridit). The ridit calculated for the jth severity code is the proportion of all individuals from the reference group falling at or below the midpoint of the jth severity code. The reference group can be a composite of all the dose groups (study 2) or the control group(study 1).

A dose level's MEAN RIDIT is an estimate of the probability that a randomly selected individual from it has a value on the underlying continuous variable greater than or equal to the value for a randomly selected individual from the reference group (Fleiss, 1986, p.81). It is an estimate of the chance that an individual in a given dose group is "worse off" than an individual in the reference group. Mean ridits for g dose groups are calculated as weighted averages of the ridits for the m individual categories.

The pairwise comparison of mean RIDITs of control and each dose group is evaluated by a chi-square statistic modified for the number of ties associated with each severity code.

Analysis 1.

In this analysis the reference group was taken to be the untreated control group one. The mean ridit rbari is an indication that a randomly selected subject from group I will have a more extreme value (greater severity) than a randomly selected subject from the reference group. SAS output for this study is presented in Table 1.

Table 1

GLOMERULONEPHROPATHY SEVERITY ANALYSIS -PESTICIDE X

Dose(ppm)

SEVERITY	GRP1	GRP2	GRP3	GRP4	ONE	RIDIT
	- number of individuals with -					
none	5	5	7	4	5	0.03472
minimum	20	25	16	13	20	0.20833
mild	21	13	18	13	21	0.49306
moderate	16	14	12	14	16	0.75000
severe	10	15	19	28	10	0.93056
	72	72	72	72	72	

RBAR1	RBAR2	RBAR3	RBAR4	POPRIDIT
0.5	0.50347	0.54350	0.63628	0.54581

STD1	STD2	STD3	STD4
0.034021	0.034021	0.034021	0.034021

95% Confidence Intervals on Individual Mean Ridits

Group	Rbar	ROUGH	USUAL	Odds
1	0.500	0.432, 0.568	0.384, 0.616	1.00:1
2	0.503	0.435, 0.572	0.387, 0.620	1.01:1
3	0.543	0.475, 0.612	0.428, 0.659	1.19:1
4	0.636	0.568, 0.704	0.524, 0.748	1.75:1

Scheffe` Analysis

Test of Equal Mean Ridits

ChiSquare	Degrees of Freedom	p_value
10.9712	3	0.011882

Group1 vs Group2

Scheffe` ChiSquare	Degrees of Freedom	p_value
.0054923	3	0.99989

Group1 vs Group3

Scheffe` ChiSquare	Degrees of Freedom	p_value
0.85590	3	0.83605

Group1 vs Group4

Scheffe` ChiSquare	Degrees of Freedom	p_value
8.41275	3	0.038209

95% Simultaneous Confidence Intervals on Mean Ridits

Group	Bonferonni	Scheffe`	Odds
2	0.388, 0.619	0.054, 0.953	1.01:1
3	0.428, 0.659	0.094, 0.993	1.19:1
4	0.521, 0.751	0.187, 1.086	1.75:1

The mean ridit for a group is the probability that a randomly selected a from individual from it has a value indicating greater severity or seriousness than a randomly selected individual from the standard group. The test for equal mean ridits is significant (p=.011882).

Rbar1=.500 -> a randomly chosen subject in the control group of non dosed subjects in group 1 is equally likely to be in a more severe injury category than the reference group which in this instance is control group 1. Rbar2=.50357 -> the subjects in dose group 2 are slightly more likely to be in a more severe injury category than the reference group of untreated subjects. Similarly rbar3=.54350 -> the subjects in dose group 3 are more likely to be in a more severe injury category . Finally, rbar4 = .63628 -> the subjects in dose group 4 are more likely to be in a more severe injury category. Looking at the population as a whole, a randomly selected subject will be in a more severe injury category compared to the control group 1 with probability of .54581 (the population mean ridit).

Hypothesis testing on differences between dosed groups and control is not significant for the dose groups 2 and 3. Dose group 4 however is statistically significantly different from the control group at the 5% level.

The odds of a randomly chosen subject in group I (i>1) of being worse off than a randomly chosen subject in reference group one are

Group	Odds
2	1.01:1
3	1.19:1
4	1.75:1

The evaluation of severity of glomerulonephrophy in male rat kidneys with dose increments of the pesticide X resulted in a statistically significant difference in the comparison of the controls and the highest dose group (1250 ppm) at the 05% level. The actual p-value was 0.038209.

Analysis 2.

Analysis 2 combines the groups within each of the severity categories rather than using only the control group as the reference group. Each method yields different results the most obvious is that now the population mean ridit is .5 rather than control group 1. SAS output for this study is presented in Table 2.

Table 2

MONONUCLEAR CELL LEUKEMIA ANALYSIS - PESTICIDE Y
Dose(mg/kg)

SEVERITY	GRP1	GRP2	GRP3	ONE	RIDIT
None	39	30	29	98	0.32667
Mild	4	5	2	11	0.69000
Moderate	2	5	9	16	0.78000
Severe	5	10	10	25	0.91667
	====	====	====	===	
	50	50	50	150	

RBAR1	RBAR2	RBAR3	POPRIDIT
0.43287	0.52633	0.5408	0.5

STD1	STD2	STD3
0.040825	0.040825	0.040825

95% Confidence Intervals on Individual Mean Ridits

Group	Rbar	ROUGH	USUAL	Odds
1	0.433	0.351, 0.515	0.294, 0.572	1.00:1
2	0.526	0.445, 0.608	0.387, 0.666	1.46:1
3	0.541	0.459, 0.622	0.401, 0.680	1.55:1

Scheffe` Analysis
Test of Equal Mean Ridits

ChiSquare	Degrees of Freedom	p_value
5.72328	2	0.057175

Group1 vs Group2

Scheffe` ChiSquare	Degrees of Freedom	p_value
3.88969	2	0.14301

Group1 vs Group3

Scheffe` ChiSquare	Degrees of Freedom	p_value
5.08331	2	0.078736

95% Simultaneous Confidence Intervals on Mean Ridits

Group	Bonferonni	Scheffe`	Odds
2	0.464, 0.723	0.168, 1.019	1.46:1
3	0.479, 0.737	0.182, 1.034	1.55:1

The mean ridit (rbar1) of .43287 is an indication of the probability that a randomly selected individual from group 1 (the control group) will have a more extreme value than a randomly selected individual from the reference group. Since this number is less than .5, we would infer that its individuals tend to have less extreme values than the subjects of the reference group. The test for equal mean ridits is significant at the 10% level (p=.057175).

Rbar2=.52633 -> the subjects in dose group 2 are slightly more likely to be in a more severe injury category than the reference group of untreated subjects. Similarly, rbar3=.5408 -> the subjects in dose group 3 are more likely to be in a more severe injury category. The mean ridit of the population is .5 which is reasonable since the reference group of untreated subjects were comprised of subjects over all the treatment groups.

Hypothesis testing on differences between dosed groups and control is not significant for the dose group 2. Dose group 3 however is statistically significantly different from the control group at the 10% level (p=.0788736).

The odds of a randomly chosen subject in group I (i>1) of being worse off than a randomly chosen subject in reference group one are

Group	Odds
2	1.46:1
3	1.55:1

The evaluation of severity of mononuclear cell leukemia in male rats with dose increments of pesticide Y (Data Set # 2) resulted in a statistically significant difference in the comparison of the controls and the highest dose group at the 10% level. The actual p_value was .078736 insignificant at the 5% level.

5.Discussion.

The statistical technique of ridit analysis was studied for its feasibility for use in environmental toxicity studies. The need for an alternative to the t-test and chi-square families of techniques arises because an ordered series of categories frequently cannot be laid out on an equal interval scale (Kantor, 1967, page 610). Severity analyses are often of this type.

A ridit is a numerical quantity which under appropriate conditions can serve as a measure of degree of injury. If ridits are being used solely for qualitative conclusions (such as would be obtained from tests of significance) one need not worry a great deal about the peculiarities of subjective scales. On the other hand, if one wants to derive quantitative results, the arithmetic mean of the ridits by severity category needs to be interpreted correctly.

For the identified distribution the mechanics of ridit analysis imposes a strong restriction (Bross, 1958, page 36). No matter what the nature of the original observations may be the mean ridits are usually uniformly distributed (the "rectangular distribution") with a uniform variance of 1/12. The sole exception occurs when nearly all of the observations fall into one or two categories(in which case a correction is needed to reduce the variance and the approximation is poorer). When the identified distribution departs from the true distribution, the true distribution will be a slightly distorted version of the rectangular distribution provided the reference set is moderately large.

Some classifications present particular difficulty. If the degrees of injury to subjects are recorded as slight, moderate, severe, disabling and

fatal, there seems no entirely satisfactory way of placing the last two classes on the same scale as the first three.

The ridit transformation is an alternative to the probit transformation being based on an empirical rather than a theoretical distribution. The usual use of the two is different. The ridit is applied in cases where there are at least three ordered response categories, while the most frequent use of probits or logits is for a dichotomous response. Ridit analysis is primarily a test of differences in location. The application of probits is an analysis of a trend in proportions.

The choice to use ridit analysis requires knowledge of the subject matter, precise information about the data, a background in statistical methods and skill in the practice of statistics.

A graphical analysis of these data sets is included as figures 1 and 2.

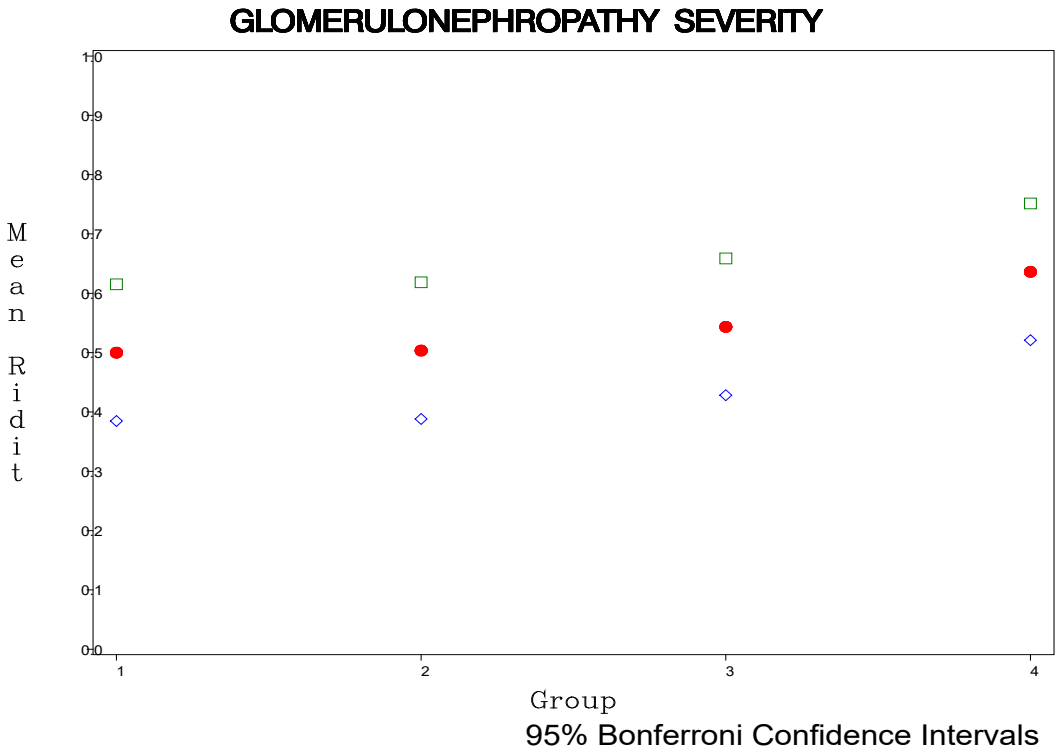

Figure 1

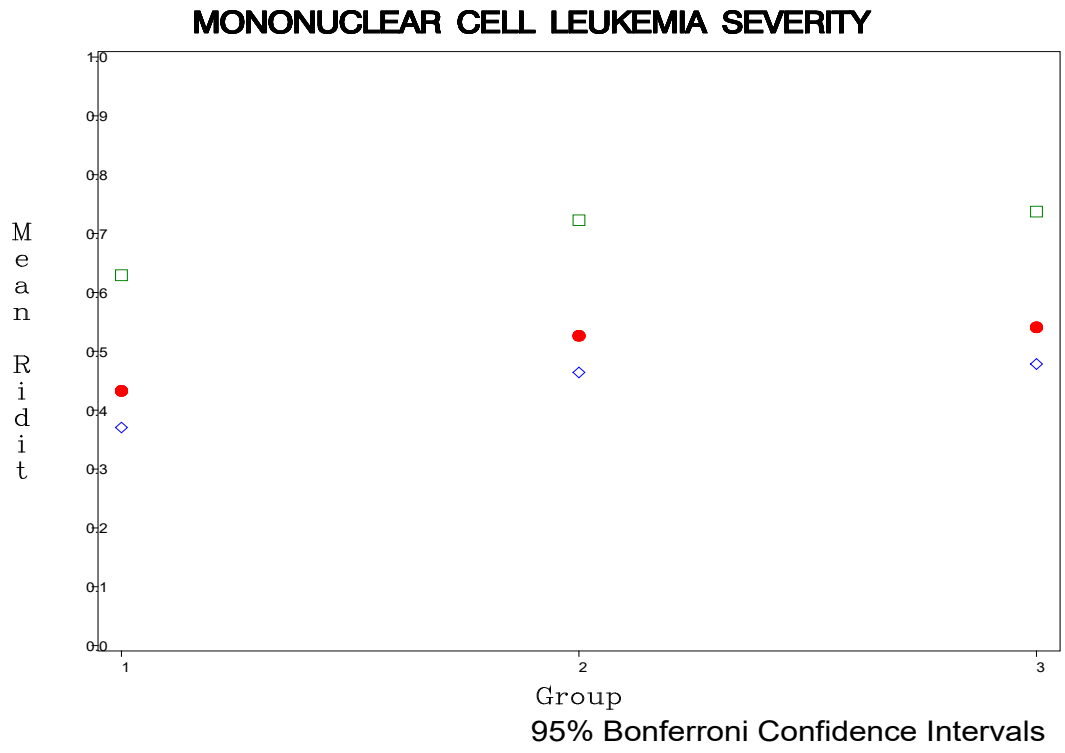

Figure 2

6. Acknowledgements.

Although the research described in this article has been funded wholly by the U.S. Environmental Protection Agency, it has not been subjected to Agency review. Therefore, it does not necessarily reflect the views of the Agency.

SAS is a registered trademark of SAS Institute Inc. in the USA and other countries. Analyses and development were done using Windows 3.1 and Windows 95 operating systems and SAS versions 6.11, 6.12. Code is operable in version 8.1 on a Windows 98 platform.

Author Contact: Mary A. Marion, 934 Marsac Street, Charlottesville, Virginia e-mail: mmstat7@gmail.com. Voice: 703-851-0446.

7. Bibliography.

Bross, Irwin D.J. (1958), How to Use Ridit Analysis, Biometrics Number 1, Volume 14, pages 18-38.

Bross, Irwin D.J. (1978), Ridit Analysis, American Journal of Epidemiology v107, pages 263-264.

Bross, Irwin D. J. (1956), Ridit Analysis of Automotive Crash Injuries, Cornell University Medical College, Department of Public Health and Preventive Medicine, Division of Automotive Crash Injury Research.

Bross, Irwin D. J. (1979), Ridit Analysis and Ranking Procedures: Reply on Biases in Judging Statistical Methods, Amerian Journal of Epidemiology, v109, pages 30-32.

Chilton, Neal W. (1982), Design and Analysis in Dental and Oral Research, 2nd Edition, Praeger Publishing Company: New York.

Cochran, William G. (19xx), Statistical Methods -Sixth Edition, The Iowa State University Press: Ames, Iowa, page 246.

Fleiss, Joseph L.(1986), The Design and Analysis of Clinical Experiments, pages 76-90, John Wiley and Sons: New York, New York.

Fleiss, Joseph L.(1983), Statistical Methods for Rates and Proportions, pages 92-108, John Wiley and Sons: New York, New York.

Fleiss, Joseph L., Chilton, Neal W. and Wallenstein, Sylvan (Nov 1979), Ridit Analysis in Dental Clinical Studies, Journal of Dental Research Volume 58, pages 2080-2084.

Flora Jr., Jairius D. (1974), A Note on Ridit Analysis - Biostatistics Technical Report Series No. 3, Department of Biostatistics, School of Public Health and Highway Safety Research Institute, University of Michigan: Ann Arbor, Michigan 48104, U.S.A.

James, Robert C. & James, Glenn (1968), Mathematics Dictionary - Third Edition, page 268, Van Nostrand Reinhold Company: New York, New York.

Kantor, S, Winklestein, W. And Ibrahim, M.A. (1968), A Note on the Interpretation of the Ridit as a Quantile Rank, American

Journal of Epidemeiology -Volume 87, pages 609-615.

Kotz, Samuel, Johnson, Norman L. and Read, Campbell B. (1982), Encyclopedia of Statistical Science, V8, pages 136-139, John Wiley and Sons: New York, New York.

Mahler, David B., Terkla, Louis G. and Van Eysden, Jan (1973), Journal of Dental Research, July-August.

Mahler, David B., Terkla, Louis G., Van Eysden, Jan and Reisbick, Morris H. (1970), Marginal Fracture vs Mechanical Properties of Amalgam, Journal of Dental Research 49: 1452-1457..

Mantel, Nathan (1979), Ridit Analysis and Related Ranking Procedures -Use at Your Own Risk, American Journal of Epidemiology published by The Johns Hopkins University School of Hygiene and Public Health.

Marion, Mary (1996), Severity Analysis Using Ridits, Proceedings of the Twenty-First Annual SAS Users Group International Conference, March 10-13, pages 957-969.

Mendenhall, William, Wackerly, Dennis D. and Scheaffer, Richard L. (1990), Mathematical Statistics with Applications -Fourth Edition, PWS-Kent Publishing Company: Boston, Massachusetts, pages 624-626.

SAS Language: Reference, Version 6, First Edition, SAS Institute Inc.: SAS Campus Drive, Cary, North Carolina, page 491.

Scott, R.E., Flora, J.D., and Marsh, J.C. (1976), An Evaluation of the 1974 and 1975 Restraint Systems, Highway Safety Research Institute, University of Michigan: Ann Arbor, Michigan.

Scheffe`, H. (1959), Analysis of Variance, John Wiley & Sons, New York, New York, pages 66-67.

Selvin, Steve (1977), A Further Note on the Interpretation of Ridit Analysis, American Journal of Epidemiology -Volume 105, pages 16-20.

Spiegel, Murray R.(1961), Statistics -Schaums's Outline Series, McGraw-Hill Book Company, page 143.

Spitzer, Robert L., Fleiss, Joseph, Kernohan, William, Lee, Joan C. and Baldwin, Ingram T. (1965), Mental Status Schedule:

Comparing Kentucky and New York Schizophrenics, Archives of General Psychiatry, Volume 12, pages 448-455.

Wynder, Ernest L., Bross, Irwin J. And Hirayama, Takeshi (1960), A Study of the Epidemiology of Cancer of the Breast, Cancer Volume 13, pages 559-601.

8. SAS Code.

The SAS code developed to analyze the above data sets is below. Both call the macro ridits for the ridit analysis. Inputs to the macro are explained in the beginning of the macro itself.

Data set #1.

```
%let mtitle=
     %str(GLOMERULONEPHROPATHY SEVERITY ANALYSIS -PESTICIDE X));
title1 &mtitle;
%let nogrp=4;
%let codeno=5;
%let compgrp=2;
%let one=sum(of Grp1-Grp&nogrp);

data d;
input severity $ Grp1-Grp&nogrp  @@;
one=&one;
cards;
none      05  05  07  04
minimum   20  25  16  13
mild      21  13  18  13
moderate  16  14  12  14
severe    10  15  19  28
;

%ridits(d,&nogrp,&codeno,&mtitle,&one);
```

Data set #2

```
%let mtitle=
     %str(MONONUCLEAR CELL LEUKEMIA SEVERITY ANALYSIS -Data Set
2);
title1 &mtitle;
%let nogrp=3;
%let codeno=4;
%let compgrp=2;
%let one=sum(of Grp1-Grp&nogrp);

data d;
input severity $ Grp1-Grp&nogrp  @@;
one=&one;
cards;
```

```
code0    39 30 29
code1    04 05 02
code2    02 05 09
code3    05 10 10
;

%ridits(d,&nogrp,&codeno,&mtitle,&one);
```

```
%macro compGroup(compGroup);
/* TIE Group and Fij Statistic Calculations */
/* Fleiss, Design and Analysis of Clinical Experiments
           page 77 (3.37) */
data fij; set ridit;
keep dum tie ndot fnum1 fdenom1 sum1-sum&noGroup
     ndot MnRidit1-MnRidit&noGroup;
tie=Group1+Group&compGroup;
ndot=sum1+sum&compGroup;
fnum1=tie*(tie-1)*(tie+1);
fdenom1=ndot*(ndot-1)*(ndot+1);
%mmeans2(fij,fnum1,sum);
data fij; merge fij meansout; by dum; drop sum;
fnum=sum;
%mmeans2(fij,fdenom1,max);
data fij; merge fij meansout; by dum; drop max;
fdenom=max;
Fij=1-(fnum/fdenom);

data scheffe; set fij; format Chi 10.5;
if _N_ > 1 then delete;
/* Fleiss, Design and Analysis of Clinical Experiments
           page 82 (3.43) */
above=12 * sum1 * sum&compGroup * (MnRidit1-MnRidit&compGroup)**2;
below = ( sum1 + sum&compGroup + 1) * Fij;
Chi=above/below;
df=&noGroup-1;
p_value=1-probchi(Chi,df);

/* OUTPUT of Scheffe ChiSquare */
title;
proc print data=scheffe split='*' noobs; var Chi df p_value;
label Chi="Scheffe`*ChiSquare" df='Degrees*of*Freedom'
p_value='p_value' ;
title "Group1 vs Group&compGroup"; run;
title;
%mend compGroup;


%macro dosums(codeno);
%local i ii stop;
data td; set td; keep sum1-sum&codeno;
sum1=0;
```

```
sum2=sum(col1);
%let stop=&codeno+1;
%do i=3 %to &stop;
   %do ii=&i-2 %to &i-2; %end;
   sum&i=sum(of col1-col&ii);
%end;
%mend dosums;


%macro equalmns(noGroup);
/* TIE Group and F Statistic Calculations */
/* Fleiss, Design and Analysis of Clinical Experiments
      page 77 (3.37) and page 82 (3.42) */
data f; set ridit;
keep dum tie ndot fnum1 fdenom1 sum1-sum&noGroup
     ndot MnRidit1-MnRidit&noGroup PopulationRidit tot;
tie=sum(of Group1-Group&noGroup);
fnum1=tie*(tie-1)*(tie+1);
fdenom1=ndot*(ndot-1)*(ndot+1);
%do j=1 %to &noGroup;
comp&j = sum&j * (MnRidit&j-PopulationRidit)**2;
%end;
tot=sum(of comp1-comp&noGroup);

%mmeans2(f,fnum1,sum);
data f; merge f meansout; by dum; drop sum;
fnum=sum;

%mmeans2(f,fdenom1,max);
data f; merge f meansout; by dum; drop max;
fdenom=max;
F=1-(fnum/fdenom);

data equalmns; set f; format Chi 10.5;
if _N_ > 1 then delete;
/* Fleiss, Design and Analysis of Clinical Experiments page 82 (3.42)
*/
above=12 * ndot * tot;
below = ( ndot + 1) * F;
Chi=above/below;
df=&noGroup-1;
p_value=1-probchi(Chi,df);
run; quit;

/* OUTPUT of Scheffe` Test of Equal Means ChiSquare */
```

```
proc print data= equalmns split='*' noobs; var Chi df p_value;
   label Chi="ChiSquare" df='Degrees*of*Freedom' p_value='p_value' ;
   title1 "Scheffe` Analysis";
   title2 "Test of Equal Mean Ridits"; run; quit; title;
%mend equalmns;


%macro grafit(noGroup,alpha);
data grafit; set interval; keep MnRiditL MnRidit MnRiditU group;
%local i;
%do i=1 %to &noGroup;
/*  MnRiditL=usualL&i; MnRiditL=SchefL&i; */
/*  MnRiditU=usualU&i; MnRiditU=SchefU&i; */
    MnRiditL=BonL&i;
    MnRidit=MnRidit&i;
    MnRiditU=BonU&i;
    Group=&i;
    output;
%end;
data grafit; set grafit;
   confid1=1-&alpha;
   confid2=put(confid1,3.2);
   confid3=substr(confid2,2);
   call symput('conlim',trim(confid3));
   run;
proc print data=grafit noobs;
  title "grafit";  run cancel;

title1 h=1.65 f='Arial' "&mtitle";
proc sgplot data=grafit;
series x=group y=MnRiditL / markers;
series x=group y=MnRidit / markers;
series x=group y=MnRiditU / markers;
xaxis grid;
yaxis grid;
run; quit;
%mend grafit;

%macro group;
%local j;
%do j=1 %to &noGroup;
%mmeans2(ridit,Group&j,sum);
   data ridit; merge ridit meansout; by dum;
   data ridit; set ridit; drop sum;
   sum&j=sum;
```

```
   product&j=Group&j*ridit;
%mmeans2(ridit,product&j,sum);
   data ridit; merge ridit meansout; by dum;
   data ridit; set ridit; drop sum;
   sumprod&j=sum;
   MnRidit&j=sumprod&j/sum&j;
   Std&j=1/(2*sqrt(3*sum&j));
   wtMnRidit&j=sum&j*MnRidit&j;
%end;
data ridit; set ridit;
  ndot=sum(of sum1-sum&noGroup);
  PopulationRidit=sum(of wtMnRidit1-wtMnRidit&noGroup)/ndot;
%mend group;


%macro interval(noGroup,alpha);
data interval; set ridit;
%local i;
quanparm=1-(&alpha/2);
quantile=probit(quanparm);
%do i=1 %to &noGroup;
   diff&i   = 1-MnRidit&i;
   sqval1&i = (MnRidit&i * diff&i) / ( sum&i-1 );
   sqval&i  = sqrt(sqval1&i);
   usual&i  = quantile * sqval&i;
   usualL&i = MnRidit&i - usual&i;
   usualU&i = MnRidit&i + usual&i;
   roughL&i = MnRidit&i - 1 / sqrt(3*sum&i);
   roughU&i = MnRidit&i + 1 / sqrt(3*sum&i);
   odds&i   = ( 0.5 + ( MnRidit&i-MnRidit1 ) ) /
        ( 1 - ( 0.5 + (MnRidit&i-MnRidit1) ) );
%end;

data interval; set interval; if _n_ > 1 then delete;
keep MnRidit1-MnRidit&noGroup diff1-diff&noGroup
     usualL1-usualL&noGroup usualU1-usualU&noGroup
     roughL1-roughL&noGroup roughU1-roughU&noGroup
     odds1-odds&noGroup quantile sum1-sum&noGroup;

data interval; set interval; drop quanparm;
%local i;
%do i=1 %to &noGroup;
   num&i = sqrt(sum&i+sum1);
   den&i = 2*sqrt(3*sum&i*sum1);
   se&i  =  num&i/den&i;
```

```
   paramter = 1- (&alpha / (2* (&noGroup-1) ) );
   B = probit(paramter);
   paramter =  &noGroup-1;
   quanparm = 1- (&alpha/2);
   S = cinv(quanparm, paramter);
   BonL&i   = MnRidit&i-MnRidit1+.5 - B*se&i;
   BonU&i   = MnRidit&i-MnRidit1+.5 + B*se&i;
   SchefL&i = MnRidit&i-MnRidit1+.5 - S*se&i;
   SchefU&i = MnRidit&i-MnRidit1+.5 + S*se&i;
%end;

data interval; set interval;
   %global conlim;
   confid1=1-&alpha;
   confid2=put(confid1,3.2);
   confid3=substr(confid2,2);
   call symput('conlim',trim(confid3));
   run;

data interval; set interval;
 file print;
    put "         "
        "    &conlim% Confidence Intervals on Individual Mean Ridits"
/;
    put "         " "Group" "   Mean Ridit"
                     "        ROUGH     " "      USUAL"
        "      ODDS" / ;
    %do i= 1 %to &noGroup;
    ii=&i;
    put "         " ii 5.0
        MnRidit&i   10.3
        roughL&i 09.3 "," roughU&i 06.3
        usualL&i 09.3 "," usualU&i 06.3
        odds&i   08.2 ":1";
    %end;
 run;
%mend interval;


%macro mmeans2(dsname,varlst,stat);
proc datasets library=work nolist; delete meansout; run cancel;
proc delete data=meansout; run; quit;
proc means data=&dsname noprint;
  var &varlst;
  output out=meansout
```

```
     n=n nmiss=nmiss mean=mean Std=Std min=min max=max range=range
     sum=sum var=var uss=uss css=css Stderr=Stderr cv=cv
/*   skewness=skewness kurtosis=kurtosis sumwgt=sumwgt */
     t=t prt=prt;
run;
data meansout; set meansout;
dum=1;
keep &stat dum; run;
%mend mmeans2;


%macro myprints;
options formdlim='';
title1;
data _null_; file print; put _page_; run;
title1 &mtitle;
proc print data=ridit noobs;
var severity Group1-Group&noGroup one two three four ridit
    product1-product&noGroup sum1-sum&noGroup sumprod1-sumprod&noGroup
    MnRidit1-MnRidit&noGroup Std1-Std&noGroup wtMnRidit1-
wtMnRidit&noGroup ndot
    PopulationRidit;
    title2 "Calculations -&noGroup Groups"; run;
options formdlim='';
title1 &mtitle;
proc print data=fij; title2 'fij'; run;
proc print data=scheffe noobs; title2 "scheffe`"; run ;

options formdlim='';
title1 &mtitle;
proc print data=f noobs;        title2 'f'       ; run ;
proc print data=equalmns noobs; title2 "equalmns"; run ;
proc print data=interval noobs label;
var  MnRidit1-MnRidit&noGroup diff1-diff&noGroup
     usualL1-usualL&noGroup usualU1-usualU&noGroup
     roughL1-roughL&noGroup roughU1-roughU&noGroup
     odds1-odds&noGroup quantile sum1-sum&noGroup
     num2-num&noGroup den2-den&noGroup se2-se&noGroup;
title2 "interval"; run;
%mend myprints;


%macro ridits(DataFile,noGroup,codeno,alpha,diagnose,mtitle,one);
/*****************************************************************
******
```

```
*  RIDIT ANALYSIS written  July 22,1995 and
*
*                 modified 1996,1997,1998,2005,2023
*
*
*
*  (C) Copyright Mary A. Marion, Jan 12, 1998
*
*
*
*  Inputs:
*
*
*
*  NOGroup  = Number of groups
*
*  CODENO = No of severity codes (levels)
*
*  ALPHA  = Significance Level such as .05
*
*  DIAGNOSE = Diagnostic print indicator (Yes or No)
*
*  MTITLE = title of the experiment
*
*  DSNAME = Input data matrix of severity codes by group
*
*  ONE    = reference population
*
*
*
*  Constraints:
*
*
*
*  Scheffe'-type comparisons between groups always compare to Group1.
*
*  Thus always enter the control group as Group1 when not combining
*
*  across all groups to form the reference population.
*                                                    *
*  all groups to form the reference population
*
*
*
*  Macros called:  compGroup, dosums, equalmns, group, interval,
```

```
*
*                   mmeans2, myprints
*
*
*
*************************************************************************
*** */
data timetrak;
time1=time();

data DataFile; set &DataFile;
one=&one;
two=one/2;

/* THREE (Column 3) Computation */
proc transpose data=DataFile out=td; var one; run;
%dosums(&codeno);
proc transpose data=td out=td2; var sum1-sum&codeno; run;
data ridit; merge td2 DataFile; rename col1=three;
keep severity one two col1 dum Group1-Group&noGroup;
dum=1;
run;

/* RIDIT Calculations */
%mmeans2(ridit,one,sum);
   data ridit; merge ridit meansout; by dum;
   data ridit; set ridit; drop sum;
   sum0=sum;
four=two+three;
ridit=four/sum0;

/* GROUP CALCULATIONS */
%group;

/* OUTPUT of table of dose group X severity levels +
   ridits for the severity categories                */
data _null_; file print; put _page_; run;
options formdlim='';
options nonumber;
proc print data=ridit noobs label;
   label severity='Severity' ridit='Ridit';
   var severity Group1-Group&noGroup /* one */ ridit;
   sum Group1-Group&noGroup /* one */;
   title &mtitle; run;
```

```
/* OUTPUT of MnRidits, Population Mean Ridit and
          standard errors of MnRidits */
data riditout; set ridit; if _n_ > 1 then delete;
format Std1-Std&noGroup 7.5;
options formdlim=' ';
title;
proc print data=riditout noobs label;
  var MnRidit1-MnRidit&noGroup PopulationRidit;
  %do i=1 %to &noGroup; label MnRidit&i="Mean Ridit&i"; %end;;
  run;
proc print data=riditout noobs;
  var Std1-Std&noGroup;
  run;

/* CONFIDENCE INTERVALS on the RIDIT MEANS */
%interval(&noGroup,&alpha);

/* TESTING the HYPOTHESIS of EQUAL MEAN RIDITS */
%equalmns(&noGroup);

/* (G-1) SCHEFFE`-type GROUP COMPARISONS to the control group (Group1)
*/
     %macro generate(noGroup);
        %do i=2 %to &noGroup;
        %compGroup(&i); %end;
        %mend;
     %generate(&noGroup);

/* Table of Confidence Intervals and Odds */
data interval; set interval;
file print;
put "           &conlim% Simultaneous Confidence Intervals on Mean
Ridits" /;
put "             "
    "Group"  "       Bonferonni"  "         Scheffe`"  "      Odds";
%do i=1 %to &noGroup;
   ii=&i;
   put "             "     ii 3.0  "  "
       BonL&i   10.3  ","  BonU&i   6.3
       SchefL&i 10.3  ","  SchefU&i 6.3
       odds&i 08.2   ":1";
%end;
run;

/* Graphical Analysis */
```

```
%grafit(&noGroup,&alpha);

/* Output of Intermediate Calculations */
%if &DIAGNOSE=No %then %goto TRAK;
%else
%if &DIAGNOSE=Yes %then %do;
%myprints;
%end;

%TRAK:
title1; title2; title3;
data timetrak; set timetrak;
time2=time();
Xtime=(time2-time1)/60;
file print;
put _page_ ;
put // "Total Execution Time is " xtime 5.3 " Minutes";
run; quit;

proc datasets nolist;
delete DataFile td td2 meansout ridit riditout fij scheffe
       f equalmns interval grafit timetrak;
run; quit;
%mend ridits;
```

SAS INPUTS (RiditsCall.sas)

```
%let mtitle=%str(GLOMERULONEPHROPATHY SEVERITY);
title1 &mtitle;
%let noGroup=4;
%let codeno=5;
%let one=Group1;
data DataFile;
input severity $ Group1-Group&noGroup  @@;
one=&one;
cards;
none      05  05  07  04
minimum   20  25  16  13
mild      21  13  18  13
moderate  16  14  12  14
severe    10  15  19  28
;
%ridits(DataFile,&noGroup,&codeno,.05,No,&mtitle,&one);

%let mtitle=%str(MONONUCLEAR CELL LEUKEMIA SEVERITY);
title1 &mtitle;
%let noGroup=3;
%let codeno=4;
%let one=sum(of Group1-Group&noGroup);
data DataFile;
input severity $ Group1-Group&noGroup  @@;
one=&one;
cards;
none     39 30 29
mild     04 05 02
moderate 02 05 09
severe   05 10 10
;
%ridits(DataFile,&noGroup,&codeno,.05,No,&mtitle,&one);
```

```
%let mtitle=%str(1974 Transportation Study -Population);
title1 &mtitle;
%let noGroup=2;
%let codeno=4;
%let one=sum(of Group1-Group&noGroup);
data DataFile;
input severity $ Group1-Group&noGroup  @@;
one=&one;
cards;
none      357 417
minor     540 330
moderate   53  33
serious    35  17
;
%ridits(DataFile,&noGroup,&codeno,.05,No,&mtitle,&one);


%let mtitle=%str(1974 Transportation Study -Group1);
title1 &mtitle;
%let noGroup=2;
%let codeno=4;
%let one=Group1;
data DataFile;
input severity $ Group1-Group&noGroup  @@;
one=&one;
cards;
none      357 417
minor     540 330
moderate   53  33
serious    35  17
;
%ridits(DataFile,&noGroup,&codeno,.05, No ,&mtitle,&one);
```