# How Ordinal is Ordinal in Ordinal Data Analysis?

Tacq, Jacques
*Hogeschool-Universiteit Brussel (HUB)*
*Stormstraat 2*
*Brussels (1000), Belgium*
*E-mail: jacques.tacq@hubrussel.be*

Nassiri, Vahid
*Vrije Universiteit Brussel (VUB), Department of Mathematics*
*Pleinlaan 2*
*Brussels (1050), Belgium*
*E-mail: vnassiri@vub.ac.be*

## 1 The importance of ordinal data

The measurement level plays a very important role in the design and the execution of a research plan. There is a relationship of tension. Certain characteristics can have a measurement level of their own, but the researcher can decide otherwise.

A well-known example is given by the school results of pupils. The measurement level is in itself ordinal, because every teacher will admit that marks are relative in nature, that the distance from 5 to 6 on a maximum of 10 is not the same as from 7 to 8, and that the marking of exams is done according to a system of comparison in pairs. The analysis of school results thus implies a jump from the ordinal to the quantitative measurement level.

The reverse occurs often as well. The determination of age by means of the trichotomy young, median, old is a reduction of the quantitative to the ordinal measurement level.

**The researcher can** therefore **do two things**. Being a kind of social architect, he can use expensive and refined materials, whereas he might have reached the same results with cheaper means. On the other hand, he can prescribe materials that are too cheap and that affect the solidity of the building.

But there are some **important advices** we would like to give. **1)** We think that one should **start** a research with materials that are **as refined as possible**. For example, one should not start with the variable age as a trichotomy, because once this is done, there is no way back. If on the other hand, one starts with age in numbers of years, then this variable is always still collapsable. **2)** A second advice, in line with the first, is that one should **not create a loss of information**. This is among others done in dummy coding. For example, variable social status with ranks low, medium and high can be coded by means of two 0-1 coded dummies $D_1$ and $D_2$, where $D_1 = 0$, $D_2 = 0$ means low status, $D_1 = 1$, $D_2 = 0$ means medium status and $D_1 = 0$, $D_2 = 1$ means high status. As such a dummy coding could also be performed on the nominal variable nationality with categories Belgian, British and Irish, which is not ordinal, this procedure creates the loss of ordinal information. This is not recommendable.

As for **ordinal variables** such as social status, some **features** are very important. Next to *identifiability* of the characteristics low, medium and high, which is also a feature of nominal variables, ordinal variables do show an *order*. But unlike quantitative variables of interval measurement level such as degrees of Celsius, *the distances between the ranks can not be determined*, and unlike quantitative variables of ratio measurement level such as age, *the ratios between the ranks can not be determined.* For example, in a status scale UU,LU,UM,LM,UL,LL (where U means upper and L means lower) we can not say that the distance between lower lower class LL and upper lower class UL would be the same as the distance between upper lower class UL and lower middle class LM. Another example of this would be a scale of occupational level: we can not say that the difference in level between unskilled workers and skilled workers is the same as the difference between skilled workers and lower employees. And also the ratios can not be determined: we can not say that the ratio between the lower upper class LU and the upper lower class UL is 5/2, because that would mean that we have coded UU,LU,UM,LM,UL,LL as 6, 5, 4, 3, 2, 1, which would be to treat an ordinal scale as if it were a ratio scale.

An extreme example of such anomaly would be the use of the alphabet by numbers: $a = 1$, $b = 2$, $c = 3, \ldots, z = 26$. We would then give my colleague Vahid the score (v=22) + (a=1) + (h=8) + (i=9) + (d=4) = 44 and I myself, Jacques, would obtain the score (j=10) + (a=1) + (c=3) + (q=17) + (u=21) + (e=5) + (s=19) = 76. And the mean of our two names would be (44+76)/2 = 60. Everybody would understand that we can not calculate the mean name of the conference members like this, but that is what we are in fact doing when we treat an ordinal scale as if it were quantitative.

## 2 Review of several ordinal data analysis methods

There are many different methods developed in order to face with modeling ordinal data, like Kruskal-Wallis analysis of variance for ranks (Kruskal and Wallis (1952)), monotonic analysis of variance (Kruskal and Carmone (1968)), isotonic regression analysis (Barlow et al. (1972)), LISREL for ordinal data (Jöreskog (2005)), log-linear models for ordinal data (Goodman (1979)), polytomous Mokken scale analysis (Mokken (1971)), multidimensional unfolding techniques (MUDFOLD), see, e.g., Coombs (1964), the all Gifi system for nonlinear multivariate analysis, e.g., homals, princals, canals, etc., (Gifi (1990)), Spearman's-RHO, Kendall's-TAU (Kendall (1962)), Somers'-D, ordinal path analysis (Smith (1974)) and many other measures and methods. In this section we are going to discuss some of the most important ones.

There are ***three main approaches*** in dealing with ordinal data. The ***first*** approach is lowering the scale and treat ordinal data as nominal, which is usually done by dummy coding, the ***second*** approach is using some appropriate re-scaling methods, and then treat ordinal data as if they are quantitative. The ***third*** approach is treating ordinal data as truly ordinal.

In this paper we will make a plea for the Kendall-group of coefficients, in which the ***third*** approach is used. Recently there have been developed some promising new methods in which the third approach seems to be chosen, i.e., the logit models with ordinal response variable. Ann A. OConnell (2006) distinguishes the cumulative odds model, the continuation ratio model and the adjacent categories model. The restriction of these models is that the ordinality is only connected with the outcome variable. The measurement level of the independent variables is not discussed. ***But in most of the mentioned procedures, the second approach is used.*** In some cases this shift of measurement level is not a capriole. For example, when we use ***Spearman's RHO*** for the social status variable with ranks UU,LU,UM,LM,UL,LL and some other variable, then the numbers 6,5,4,3,2,1 are used for the six ranks, whence the differences between the ranks are assumed to be equal. The same occurs when one has Likert-scaled variables with ranks strongly agree, agree, don't know, disagree, strongly disagree (++,+,?,-,–) and one performs ***alpha factor analysis*** to calculate Cronbach's alpha. One then assigns the numbers 5,4,3,2,1 to the five ranks ++,+,?,-,–. We would like to call this the ***vulgar quantification***. It is a thoughtless and inadmissible shift of measurement level which is beyond words.

Something similar is done in *Kruskal and Wallis analysis of variance*: the dependent variable is ordinal and receives rank numbers 1,2,3,4, ect., of which arithmetic means are calculated, just like in Spearman's coefficient. And also *isotonic regression* of Iman and Conover (1979) is totally based on Spearman's RHO. *Monotonic ANOVA* of Kruskal and Carmone (1968) finds a monotonic function of the original response variable Y that is maximally correlated with an additive model. A measure of departure of fit, called stress, is minimized, subject to obeying the monotonic constraints in the criterion variable. Here too, rank numbers are used and arithmetic means thereof are calculated, soundly and thoroughly.

A way out seems to be *Jöreskog's LISREL for ordinal data.* Here each ordinal variable, say $B$, is considered to be a manifestation of an underlying unobserved metric variable $Z$, which is standard normally distributed. The ranks of $B$ then correspond to some values of $Z$, which are called thresholds and which are computed by using the inverse cumulative normal distribution. For the correlation between two ordinal variables $A$ and $B$, one assumes underlying variable $Z$ for $B$ and underlying variable $Y$ for $A$, and the correlation between $A$ and $B$ is called polychoric correlation. An underlying regression model is assumed and an iterative procedure minimizes the distance between the observed frequencies and the expected frequencies as given by the underlying variable model. Often a two step procedure is applied: 1) estimation of the thresholds; 2) iterative estimation of the polychoric correlation. An extension of this procedure is an approach based on the matrix of polychoric correlations between many ordinal variables in a multivariate analysis.

It goes without saying that this LISREL for ordinal variables is totally dependent on the assumption of underlying unobserved metric variables that are standard normally distributed. Kampen (2001) points out that researchers in practice are unable to verify the assumptions about underlying variables, and I add, let alone the existence of such underlying variables. It looks like as if we are in the cave of Plato, with a beautiful reality outside with a shining sun and green grass, which we will never see, and where we have to rely on the shadow-images on the walls of the cave. So, this kind of ordinality of Jöreskog's LISREL could be called "**Plato-ordinality**" or "**Cave-ordinality**". It is highly imaginary and tends to idealism rather than to a realistic approach.

> In Plato's Allegory of the Cave, Socrates describes a group of people who have lived chained to the wall of a cave all of their lives, facing a blank wall. The people watch shadows projected on the wall by things passing in front of a fire behind them, and begin to ascribe forms to these shadows. According to Socrates, the shadows are as close as the prisoners get to viewing reality. He then explains how the philosopher is like a prisoner who is freed from the cave and comes to understand that the shadows on the wall do not make up reality at all, as he can perceive the true form of reality rather than the mere shadows seen by the prisoners.

Another approach is the all *Gifi system* for nonlinear multivariate analysis, e.g., categorical regression analysis, categorical pca, nonlinear canonical correlation analysis, etc., which can be analyzed in SPSS-Categories. Here the ordinal variables are assigned numeric values through a process called optimal scaling. Such numeric values are referred to as category quantifications. Optimal scaling is achieved by the minimization of a least squares loss function For example, in nonlinear PCA, model estimation and optimal quantification are alternated through use of an iterative algorithm that converges to a stationary point where the optimal quantifications of the categories do not change anymore. Optimal scaling with alternating least squares replaces the category labels (here: ranks) with category quantifications in such a way that as much as possible of the variance in the quantified variables is accounted for. Specifically, the method maximizes the first $p$ eigenvalues of the correlation matrix of the quantified variables, where $p$ indicates the number of components that are chosen in the analysis. This criterion is equivalent to the previous statement that the aim of optimal scaling is to maximize the VAF (Variance Accounted For) in the quantified variables.

**In Gifi** there are three levels of analysis, nominal, ordinal and numeric. But here, in our view, ordinal does not really mean ordinal. As indicated, the category quantifications maximize the VAF. Now, in doing so the different levels of analysis point to different requirements. In the **nominal case**, the requirement is that individuals who scored the same category on the original variable also obtain the same quantified value. But notice that the choice of nominal level of analysis does not imply that the original variable is measured at the nominal level. For an ordinal or even a numeric variable, it is equally well possible to choose the nominal level of analysis. In the **ordinal case**, the additional requirement is that the quantifications of the categories should respect the order of the original variables. If the original category labels are increasing, then the transformation to quantifications will be monotonically non-decreasing, where some consecutive categories may obtain the same quantification, referred to as ties, so that a step function is obtained. If all variables are at the **numeric level** of analysis, no optimal quantification is performed and variables are simply standardized, in which case potential nonlinear relationships among variables are not accounted for. So, if one wishes to account for nonlinear relations between numeric variables, one has to choose for a non-numeric level of analysis, ordinal or nominal.

The **ordinal level of analysis** described above makes use of **step functions**. As an alternative, one can use a procedure of **smoothing**. This is done in spline transformation, monotonic or nonmonotonic. Now, suppose that the relationship of a specific variable with other variables in a multivariate analysis is such, that the quantifications of categories go first up and then down, then this will be seen only when the nominal or nonmonotonic spline level is chosen. However, the ordinal and monotonic spline level wil not show this; they will show an increase of the quantifications of categories up to a certain point and for the higher category labels there will be ties. And the numeric level of analysis will teach us nothing, the plot of original scores and quantifications after transformation will just show a linear function, a straight line.

The consequence is that ordinal is not really ordinal in this Gifi-system. First of all, as stated above, VAF is the criterion in use, which means that the behaviour of a variable is guided by its relationship with all the other variables in the analysis. And secondly, the choice of ordinal level of analysis, and also monotonic spline, brings about a smoothing out with ties, and consequently, only the nominal level of analysis, and also nonmonotonic spline, show us the real behaviour of a variable, e.g., going first up and then down. B-splines do not alter the basic argument. In a personal visit and conversation with the Gifi-people at University of Leiden in The Netherlands, they admitted that ordinal is not really ordinal. We will call this "**the brain in a vat ordinality**", because the criterion of ordinality in use is not ordinality, but something else, e.g., the relationship with other variables or the variance accounted for (VAF) combined with the procedure of smoothing out.

> In philosophy, the *brain in a vat* is an element used in a variety of thought experiments intended to draw out certain features of our ideas of knowledge, reality, truth, mind, and meaning. It is drawn from the idea, common to many science fiction stories, that a mad scientist, machine or other entity might remove a person's brain from the body, suspend it in a vat of life-sustaining liquid, and connect its neurons by wires to a supercomputer which would provide it with electrical impulses identical to those the brain normally receives. According to such stories, the computer would then be simulating reality (including appropriate responses to the brain's own output) and the person with the "disembodied" brain would continue to have perfectly normal conscious experiences without these being related to objects or events in the real world.

The mad scientist, let us call him Frankenstein, is here the Gifi-system, which is calculating at a very sophisticated level, attaching the variables at a supercomputer so to speak, and in this way simulating a world, which is in fact not equivalent to the real world.

As we are in the dark about the real world, we might also call this "**the Zhuangzi ordinality**" or "**butterfly ordinality**", referring to the butterfly dream of Zhuangzi.

> Once Zhuangzi dreamt he was a butterfly, a butterfly flitting and fluttering around, happy with himself and doing as he pleased. He didn't know he was Zhuangzi. Suddenly he woke up and there he was, solid and unmistakable Zhuangzi. But he didn't know if he was Zhuangzi who had dreamt he was a butterfly, or a butterfly dreaming he was Zhuangzi. Between Zhuangzi and a butterfly there must be some distinction!

## 3 Treating ordinal data as truly ordinal. The case of Kendall's TAU

We know that it will sound old-fashioned to dig up the treatment of Sir Maurice Kendall of the forties. And that is what many people said at first, when they heard of our intention. But on second thought most colleagues agree that the work of Kendall and his successors must be seen as a forgotten treasure. The reason for this is clearly formulated by Jarl Kampen (2001) in his dissertation on ordinal variables and their analysis. In a chapter on measures based on concordancy, he deals with Kendall's TAU, Somers' D and Goodman and Kruskal's GAMMA and he writes: "These measures have gradually lost their appeal in applied research, because their applicability in multivariate research is limited. For example, an alternative measure is the polychoric correlation , which can be used in multivariate regression and factor analytic models. Equivalent extensions have never been developed for either Kendall's $\tau_b$, Goodman and Kruskal's GAMMA or Somers' D, nor are they likely to be developed because of computational difficulties in determining standard errors of the involved parameters and because several alternative models are already in high degree of development."

We would like to protest here. For, there exists a literature about multiple tau, partial tau and partial Somers' d, which is on a parallel with multiple correlation, partial correlation and partial regression coefficient, respectively. In the limited space we have here, we would like to raise a corner of the veil. Of course, Jarl Kampen is right in his observation that several alternative models are in high degree of development. LISREL for ordinal data, loglinear modeling for ordinal data, the all Gifi system for nonlinear multivariate analysis, and others, have overruled the Kendall-group of coefficients, whence the latter have been snowed under. But then we would like to protest against these historical developments.

One of the reasons of our protest is the fact that, unlike many developments in the analysis of ordinal data, the *Kendall-group* of coefficients are *truly ordinal.* To explain this, we compare with Spearman's coefficient. Spearman's coefficient has an enormous drawback : the shift of measurement level. Variables measured at the ordinal level are treated as if they were quantitative, with equal intervals between the distinctions. Take, for example, occupational level in which the lowest dictinction is the unskilled worker, followed by the skilled worker and thereafter by the lower employee etc. Then it goes without saying that we are in the dark about the distances between unskilled and skilled worker and between skilled worker and lower employee. We can not treat these distances as equal, and that is what Spearman's coefficient does, because the ranks are given the numbers 1, 2, 3 etc. This will not be done by Kendall's coefficient TAU and that is why TAU will be really superior in terms of appropriateness for the ordinal measurement level.

Based on Daniels (1944) Kendall's TAU is equal to Pearson's r in which "variation" (sum of squares) and "covariation" (sum of cross-products) are replaced with "ordinal variation" and "ordinal covariation". The generalized coefficient of covariation $\Gamma$ is as follows,

$$\Gamma = \frac{\sum x_{ij} y_{ij}}{\sqrt{\sum x_{ij}^2 \sum y_{ij}^2}}$$

Consider $X$ and $Y$ as interval measures and $x_{ij} = X_i - X_j$ and $y_{ij} = Y_i - Y_j$, then $\Gamma$ is Pearson's r, if $X$ and $Y$ are rankings and $x_{ij}$ and $y_{ij}$ are the differences between ranks, then $\Gamma$ is Spearman's RHO. Also, if we assign $+1$ to $x_{ij}$ if $X_i$ is greater than $X_j$ and $-1$ otherwise and the same for $y_{ij}$, then $\Gamma$ will be Kendall's TAU. Somer's D which is calculated using Kendall's TAU has also an analogy with regression coefficients.

There is an interesting connection between how Kendall's TAU treats ordinality and French sociologist Pierre Bourdieu's analysis of class structure in society. In his book 'Distinction', on the basis of thorough survey-analysis and using criteria such as economic, social and cultural capital, Bourdieu distinguishes three classes: dominant class, middle class and working class, and also class fractions within these classes. It is very important to keep in mind that this stratification of society does not lend itself to a substantialist reading. Bourdieu expresses that the distinctions between classes get their meaning by means of a relational logic. The manners and life-styles to which distinctions refer do not have an innate quality giving them the meaning they have, but emerge by virtue of differences in social space, drawing difference out of the undifferentiated, so to speak.

The analogy with ordinal measurement is straightforward. Making use of +1 when a difference $(x_i - x_j)$ shows $x_i > x_j$, and -1 when $x_i < x_j$ is the case, and 0 when $x_i = x_j$, and doing the same for $Y$, no substantialist reading of ordinality but rather a relational logic - we might call it a relational topology - is involved in Kendall's TAU. In other words, Kendall's TAU measures the association between variables without measuring the variables themselves. It is therefore a pity (a shame?) that Kendall's and Somers' coefficients and their multivariate extensions have not received a more important status in the statistical literature. We would like to invite statisticians in the world to forget about Spearman, forget about LISREL for ordinal data, forget about the all Gifi treatment of ordinality, and **put the forgotten treasure, which Kendall's TAU is, again in the forefront**.

## 4 Conclusion

The $18^{th}$ century philosopher Immanuel Kant has written three Critiques, Critique of Pure Reason, Critique of Practical Reason and Critique of Judgement. The latter contains two parts, one on aesthetic and one on teleology. In the part on teleology Kant asks the question whether there are ends in nature. Does the giraffe have a long neck in order to reach food at a certain height? Are beast of prey eyes on the wings of butterflies present for hostile birds to be deterred? Do flies have compound eyes in order to be quick in escaping from enemies? Does the skin of the chameleon take the colour of the environment in order to hide from possible enemies? In other words, are there ends in nature? Kant's answer is: No, there are not. But he gives the advice to scientists to do as if there are ends, because the teleological form of questioning is a stimulation for scientific research and it is a good attitude to suppose that nothing exists just like that, without an end.

This as if philosophy, which is even the title of a voluminous book of Hans Vaihinger 'Die Philosophie des Als Ob' (The Philosophy of the As If), seems to be the attitude of many statisticians dealing with ordinal data. They do as if ordinal variables are scaled on the quantitative measurement level, even if they in fact know better. This attitude looks very much as if a mechanism of Suspension of Disbelief is into play, which is operating when we read a book or look a movie and we are convinced that the fantastic world which is created can not be true, but where we prolonge our belief in the story, not because we really believe it but because we like this world of fantasy so much.

> **Suspension** of **disbelief** is a formula for justifying the use of fantastic or non-realistic elements in literature. It was put forth in English by the poet and aesthetic philosopher Samuel Taylor Coleridge, who suggested that if a writer could infuse a "human interest and a semblance of truth" into a fantastic tale, the reader would suspend judgement concerning the implausibility of the narrative. The phrase "**suspension of disbelief**" came to be used more loosely in the later 20th century, often used to imply that the onus was on the reader, rather than the writer, to achieve it. It might be used to refer to the willingness of the audience to overlook the limitations of a medium, so that these do not interfere with the acceptance of those premises. These fictional premises may also lend to the engagement of the mind and perhaps proposition of thoughts, ideas, art and theories.

In this contribution we have argued against such an attitude. We have chosen for an old contribution of the forties, which is Kendall's TAU and the coefficients related to it. There is no as if philosophy or suspension of disbelief here. On the contrary, ordinality is treated as true ordinality. And it has even the bonus that scientific thinking of ordinality in terms of a relational logic is contained in it.

## REFERENCES (RÉFERENCES)

[1] Barlow, Bartholomew, Bremner and Brunk, *Statistical Inference Under Oreder Restrections: The Theory and Application of Isotonic Regression*, John Wiley and Sons, 1972, New York.

[2] Clogg, *Statistical Models for Ordinal Variables*, Sage, 1994, London.

[3] Coombs, *A Theory of Data*, John Wiley and Sons, 1964, New York.

[4] Daniels, 'The Relation between Measures of Correlation in the Universe of Sample Permutations', *Biometrika*, **33**, 1944, pp. 129-135.

[5] Gifi, *Nonlinear Multivariate Analysis*, John Wiley and Sons, 1990, New York.

[6] Goodman, 'Simple models for the analysis of association in cross-classifications having ordered categories', *Journal of American Statistical Association (JASA)*, **76**, 1979, pp. 537- 552.

[7] Hawkes, 'The multivariate analysis of ordinal measures', *American Journal of Sociology (AJS)*, **76**, 1971, pp. 908-926.

[8] Jöreskog, 'Structural Equation Modeling with ordinal variables using LISREL', 2005, `http://www.ssicentral.com/lisrel/techdocs.ordinal.pdf`.

[9] Kendall, *Rank Correlation Methods*, Hafner, 1962, New York.

[10] Kruskal, 'Ordinal measures of association', *Journal of American Statistical Association (JASA)*, **53**, 1958, pp. 814 - 861.

[11] Kruskal and Carmone, *Use and Theory of MONANOVA*, Unpublished paper, 1968, Bell Laboratories.

[12] Kruskal and Wallis, 'Use of ranks in non-criterion variance analysis', *Journal of American Statistical Association (JASA)*, **47**, 1952, pp. 583- 621.

[13] Mokken, *A Theory and Procedure of Scale Analysis*, De Gruyter, 1971, Berlin.

[14] O'Connell, 'Logistic Regression Models for Ordinal Response Variables, *Sage*, Sage, London, 2006.

[15] Smith, 'Continuities in ordinal path analysis, *Social Forces*, **53**, 1974, pp. 200 - 253.

[16] Iman and Conover, 'The use of the rank transform in regression, *Technometrics*, **21**, 1979, pp. 499-509.

[17] Kampen, '*The Adequacy and Interpretation of Models for Ordinal Association*, 'PhD Dissertation submitted at Catholic University of Brussels', Brussels, 2001.