COEFFICIENTS OF MULTIPLE ASSOCIATION FOR RANKED DATA

Alan Agresti, University of Florida

1. INTRODUCTION

In this article, we propose a coefficient of multiple rank association $\tau_{Y\cdot\underline{X}} = \tau_{Y\cdot X^{(1)},\ldots,X^{(k)}}$ for describing a specific aspect of the association between a dependent variable Y and a set of independent variables $(X^{(1)},X^{(2)},\ldots,X^{(k)})$, when the available statistics consist of rankings on these variables for a sample of observations. In the next section $\tau_{Y\cdot\underline{X}}$ is defined as a generalization of Kendall's tau for two variables, in that it is based on the orderings for pairs of observations on each of the variables. The measure may be interpreted as a weighted average of the absolute values of the Kendall's taus between Y and each $X^{(i)}$ over sets of pairs with fixed orderings on $\{X^{(\nu)},\nu=1,2,\ldots,k\}$. In addition, $\tau_{Y\cdot\underline{X}}$ has a proportional reduction in error interpretation based on predicting pairwise ordering on Y using pairwise orderings on $\{X^{(\nu)}\}$. The extent of the increase in value of $\tau_{Y\cdot\underline{X}}$ as additional independent variables are added to the system is thus a measure of the improvement in predictive ability of these pairwise orderings on Y, for the given prediction rule. An additional coefficient $\tau^{(2)}_{Y\cdot\underline{X}}$, based on a different prediction rule, is also considered and seen to be more useful that $\tau_{Y\cdot\underline{X}}$ when $k=2$.

In Section 3, the coefficients are generalized for application when there are some tied ranks, or the variables are ordinal caregorical in nature. The measure $\tau'_{Y\cdot\underline{X}}$ is defined in terms of all pairs of observations untied with respect to Y, and is seen to have similar properties as the coefficient $\tau_{Y\cdot\underline{X}}$ defined for the full-rank (no ties) case.

Examples of the calculation of these measures are given in a corresponding technical report (see Agresti (1976)). Also, asymptotic sampling distributions are considered in the report, as well as comparisons with other ordinal measures of multiple association which have been formulated.

2. MULTIPLE TAU COEFFICIENTS

The multiple tau coefficients which we shall define in this section are based on a generalization of the proportional reduction in error interpretation for the absolute value of Kendall's tau (denoted by τ_{YX} for the population value) between two variables Y and X (see, e.g., Ploch (1974)). For this bivariate case, let C and D represent the numbers of concordant and discordant pairs of observations, and suppose that there are no tied pairs.

If one were to predict at random for each of the $n(n-1)/2$ pairs of observations whether that pair was concordant or discordant (i.e., for each pair, predict concordance with probability 1/2, predict discordance with probability 1/2), the expected number of prediction errors would be $(C+D)/2 = n(n-1)/4$. If, on the other hand, one knows that $\tau_{YX} > 0$ and predicts concordance for each pair, the number of errors would be D. This results in a proportional reduction in error of

$$\frac{n(n-1)/4 - D}{n(n-1)/4} = \frac{C-D}{n(n-1)/2} = \tau_{YX} \qquad (2.1)$$

Similarly, if one knows $\tau_{YX} < 0$ and always predicts discordance, the proportional reduction in error is

$$\frac{D-C}{n(n-1)/2} = |\tau_{YX}|.$$

One could interpret $|\tau_{YX}|$, then, as the proportional reduction in error which results from predicting the ordering of pairs of observations on Y, based on having knowledge of the orderings of the pairs on X (and using the rule whereby the majority ordering is always predicted), relative to possessing no information about the orderings on X.

2.1 Definition of $t_{Y\cdot\underline{X}}$

Now, suppose that we wish to describe the association between a dependent variable Y and a collection of independent variables $\underline{X} = (X^{(1)},\ldots,X^{(k)})$. We shall next construct a similar type of coefficient with predictions of the ordering on Y based on the orderings on the $\{X^{(\nu)}, \nu=1,\ldots,k\}$ for each pair of observations. In this section, for simplicity, we shall assume that there are no tied pairs with respect to any of the variables.

Let $(Y_1,\ldots,Y_n)$, $(X_1^{(1)},\ldots,X_n^{(1)}),\ldots,(X_1^{(k)},\ldots,X_n^{(k)})$ denote the rankings on $Y,X^{(1)},\ldots,X^{(k)}$ for n observations in some sample, and for a pair of observations (i,j), let

$$S_\nu(i,j) = S\left[(Y_j-Y_i)(X_j^{(\nu)}-X_i^{(\nu)})\right],\ \nu=1,\ldots,k \qquad (2.2)$$

where S is the sign function

$$S[u] = \begin{cases} -1, & u<0 \\ 0, & u=0 \\ 1, & u>0 \end{cases}$$

Also, denote $(S_1(i,j),\ldots,S_k(i,j))$ by $\underline{S}(i,j)$, and let

$$A(\underline{\delta}) = A(\delta_1,\ldots,\delta_k) = \{(i,j):\ \underline{S}(i,j) = \underline{\delta}\}.$$

For example, $A(1,1,\ldots,1)$ is the set of pairs of observations which are simultaneously concordant between Y and each $X^{(\nu)}$, $\nu=1,\ldots,k$. If the pair of observations (i,j) is in $A(\underline{\delta})$, then that pair is $Y-X^{(\nu)}$ concordant (discordant) if $\delta_\nu = 1$ ($\delta_\nu=-1$). Notice that the $\{A(\delta_1,\ldots,\delta_k),\ \delta_\nu=\pm 1,\ \nu=1,\ldots,k\}$ create a partition of the $n(n-1)/2$ pairs of observations. Let

$$D_k = \{(\delta_1,\ldots,\delta_k):\ \delta_\nu = \pm 1,\ \nu = 1,\ldots,k\}, \quad (2.3)$$

and denote by $N(\underline{\delta}) = N(\delta_1,\ldots,\delta_k)$ the number of pairs of observations in the set $A(\underline{\delta})$, so that

$$\sum_{D_k} N(\underline{\delta}) = n(n-1)/2$$

Now, for each element $\underline{\delta}$ of D_k, $A(\underline{\delta}) \cup A(-\underline{\delta})$ is the set of pairs of observations with a certain fixed ordering on the $\{X^{(\nu)},\ \nu=1,\ldots,k\}$, namely

$$\begin{aligned} &X^{(u)} - X^{(w)} \text{ concordant if } \delta_u\delta_w = 1\\ &X^{(u)} - X^{(w)} \text{ discordant if } \delta_u\delta_w = -1,\\ &\qquad 1 \le u \le w \le k. \end{aligned}$$

There are several ways in which one could form predictions for the Y ordering for pairs in this set. The prediction rule for the coefficient to which we shall devote primary attention specifies that for each set $A(\underline{\delta}) \cup A(-\underline{\delta})$ of pairs with fixed orderings on the $\{X^{(\nu)}\}$, one should predict ordering on Y such that

$$\begin{aligned} &\underline{S}(i,j) = \underline{\delta} \text{ if } N(\underline{\delta}) \ge N(-\underline{\delta})\\ &\underline{S}(i,j) = -\underline{\delta} \text{ if } N(\underline{\delta}) < N(-\underline{\delta}) \end{aligned} \quad (2.4)$$

That is, if the majority of pairs in $A(\underline{\delta}) \cup A(-\underline{\delta})$ are $Y-X^{(\nu)}$ concordant (discordant), predict the orderings on Y for pairs in this set corresponding to $Y-X^{(\nu)}$ concordance (discordance).

According to this prediction rule, the number of prediction errors for pairs in $A(\underline{\delta}) \cup A(-\underline{\delta})$ is $\min[N(\underline{\delta}),N(-\underline{\delta})]$. On the other hand, random predictions of $Y-X^{(\nu)}$ concordance or discordance (with probability 1/2 for each) for pairs in $A(\underline{\delta}) \cup A(-\underline{\delta})$ would correspond to an expected number of errors of $[N(\underline{\delta}) + N(-\underline{\delta})]/2$. When predictions are considered over all pairs in all such sets with fixed $\{X^{(\nu)}\}$ ordering, the proportional reduction in error obtained from utilizing knowledge of ordering on the $\{X^{(\nu)}\}$ is

$$\begin{aligned} t_{Y\cdot\underline{X}} &= \frac{\sum_{D_k}[N(\underline{\delta}) + N(-\underline{\delta})]/2 - \sum_{D_k}\min[N(\underline{\delta}),N(-\underline{\delta})]}{\sum_{D_k}[N(\underline{\delta}) + N(-\underline{\delta})]/2}\\ &= \frac{n(n-1)/2 - \sum_{D_k}\min[N(\underline{\delta}),N(-\underline{\delta})]}{n(n-1)/2}\\ &= \frac{\frac{1}{2}\sum_{D_k}|N(\underline{\delta}) - N(-\underline{\delta})|}{n(n-1)/2} \end{aligned} \quad (2.5)$$

The factor of $\frac{1}{2}$ occurs here and in some subsequent formulas due to the fact that both $|N(\underline{\delta}) - N(-\underline{\delta})|$ and $|N(-\underline{\delta}) - N(\underline{\delta})|$ occur in these sums when D_k is used as the index set.

Notice that $t_{Y\cdot\underline{X}}$ may be written as

$$t_{Y\cdot\underline{X}} = \sum_{D_k}\lambda(\underline{\delta})\ \frac{|N(\underline{\delta}) - N(-\underline{\delta})|}{N(\underline{\delta}) + N(-\underline{\delta})} \quad (2.6)$$

where $\lambda(\underline{\delta}) = \dfrac{N(\underline{\delta})}{n(n-1)/2}$ is the proportion of the $n(n-1)/2$ pairs of observations which are in $A(\underline{\delta})$. Letting

$$t(\underline{\delta}) = [N(\underline{\delta}) - N(-\underline{\delta})]/[N(\underline{\delta}) + N(-\underline{\delta})],$$

we see that

$$t_{Y\cdot\underline{X}} = \tfrac{1}{2}\sum_{D_k}(\lambda(\underline{\delta}) + \lambda(-\underline{\delta}))\,|t(\underline{\delta})| \quad (2.7)$$

is a weighted average of the absolute values of Kendall's tau type measures calculated within each set $A(\underline{\delta}) \cup A(-\underline{\delta})$ of orderings on the $\{X^{(\nu)}\}$. Since the joint orderings of the $\{X^{(\nu)}\}$ are fixed within $A(\underline{\delta}) \cup A(-\underline{\delta})$, $|t(\underline{\delta})|$ is in fact the absolute value of Kendall's tau between Y and each of the $X^{(\nu)}$ ($\nu=1,\ldots,k$), for that set of pairs.

The calculation of a coefficient such as $t_{Y\cdot\underline{X}}$ is based typically on a sample from some real or conceptual population of interest. Letting $P(\underline{\delta})$ denote the proportion of pairs of observations in $A(\underline{\delta})$ in this population, the corresponding population value of this coefficient is

$$\tau_{Y\cdot\underline{X}} = \tfrac{1}{2}\sum_{D_k}|P(\underline{\delta}) - P(-\underline{\delta})|. \quad (2.8)$$

Alternatively, let

$$D_M = \{\underline{\delta}:\ P(\underline{\delta}) > P(-\underline{\delta})\},\quad D_m = \{\underline{\delta}:\ P(\underline{\delta}) < P(-\underline{\delta})\}. \quad (2.9)$$

Then, for simplicity, we could rewrite the population coefficient as

$$\tau_{Y\cdot\underline{X}} = P_M - P_m = \sum_{D_M}[P(\underline{\delta}) - P(-\underline{\delta})], \quad (2.10)$$

where $P_M = \sum_{D_M} P(\underline{\delta}) = \Pr[S(i,j) \text{ in } D_M]$ and

$$P_m = \Pr[S(i,j) \text{ in } D_m]$$

for a randomly selected pair (i,j). We shall refer to the set of pairs indexed by D_M as those with <u>majority ordering</u> on Y with respect to $\{X^{(\nu)}\}$, and by D_m as those with <u>minority ordering</u> on Y with respect to $\{X^{(\nu)}\}$. Thus $\tau_{Y\cdot\underline{X}}$ is also similar in structure to Kendall's tau in that it may be interpreted as the difference in the probabilities of two types of pairs of observations.

2.2 Properties

We shall next consider some of the basic properties of $t_{Y\cdot\underline{X}}$. It is clear from the definition that $t_{Y\cdot\underline{X}}$ is invariant under order-preserving transformations on any of the variables. In the simple bivariate case,

$$t_{Y\cdot X} = \frac{|N(1) - N(-1)|}{n(n-1)/2} = \frac{|C - D|}{n(n-1)/2} \quad (2.11)$$

equals the absolute value of Kendall's tau between Y and X, the difference between the proportions of concordant and discordant pairs of observations. In the trivariate case,

$$t_{Y\cdot X^{(1)},X^{(2)}} = \{|N(1,1) + N(-1,-1)| + |N(1,-1) - N(-1,1)|\}/n(n-1)/2$$

$$= \frac{2}{n(n-1)}\max\{|(N(1,1) + N(1,-1)) - (N(-1,1) + N(-1,-1))|, |(N(1,1) + N(-1,1)) - (N(1,-1) + N(-1,-1))|\}$$

$$= \max\{|t_{YX^{(1)}}|, |t_{YX^{(2)}}|\} \quad (2.12)$$

The behavior of $t_{Y\cdot \underline{X}}$ becomes less trivial when the number of independent variables k exceeds two, as the simultaneous predictive power available from $(X^{(1)},\ldots,X^{(k)})$ may exceed that of the one most strongly associated with Y. In general, $t_{Y\cdot \underline{X}}$ is monotone increasing as the set of independent variables increases in size. To see that $t_{Y\cdot X^{(1)},\ldots,X^{(k)}} \le t_{Y\cdot X^{(1)},\ldots,X^{(k+1)}}$, we need only note that the partition of pairs into sets $\{A(\delta_1,\ldots,\delta_{k+1}) \cup A(-\delta_1,\ldots,-\delta_{k+1})\}$ with similar orderings on the $\{X^{(\nu)}, \nu = 1,\ldots,k+1\}$ is a subdivision of the partition $\{A(\delta_1,\ldots,\delta_k) \cup A(-\delta_1,\ldots,-\delta_k)\}$, and thus $\sum_{D_k} |N(\delta_1,\ldots,\delta_k) - N(-\delta_1,\ldots,-\delta_k)| = \sum_{D_k} |N(\delta_1,\ldots,\delta_k,1) + N(\delta_1,\ldots,\delta_k,-1) - N(-\delta_1,\ldots,-\delta_k,1) - N(-\delta_1,\ldots,-\delta_k,-1)|$

$$\le \sum_{D_k} |N(\delta_1,\ldots,\delta_k,1) - N(-\delta_1,\ldots,-\delta_k,-1)| + \sum_{D_k} |N(\delta_1,\ldots,\delta_k,-1) - N(-\delta_1,\ldots,-\delta_k,1)|$$

$$= \sum_{D_{k+1}} |N(\delta_1,\ldots,\delta_{k+1}) - N(-\delta_1,\ldots,-\delta_{k+1})|.$$

Notice that $t_{Y\cdot X^{(1)},\ldots,X^{(k)}} = t_{Y\cdot X^{(1)},\ldots,X^{(k+1)}}$ if and only if for each choice of $(\delta_1,\ldots,\delta_k)$, either

$$N(\delta_1,\ldots,\delta_k,1) \ge N(-\delta_1,\ldots,-\delta_k,-1)$$
$$\text{and } N(\delta_1,\ldots,\delta_k,-1) \ge N(-\delta_1,\ldots,-\delta_k,1),$$

or $\quad (2.13)$

$$N(\delta_1,\ldots,\delta_k,1) \le N(-\delta_1,\ldots,-\delta_k,-1)$$
$$\text{and } N(\delta_1,\ldots,\delta_k,-1) \le N(-\delta_1,\ldots,-\delta_k,1);$$

that is, if the refinement in the partition of pairs by adding $X^{(k+1)}$ to the system does not result in a change in the predictions of $Y - X^{(\nu)}$ concordance or discordance for any of the pairs, $\nu = 1,\ldots,k$. In particular, if $|t_{X^{(\ell)}X^{(k+1)}}| = 1$ for some ℓ $(1 \le \ell \le k)$, then the partition is unchanged and $t_{Y\cdot X^{(1)},\ldots,X^{(k)}} = t_{Y\cdot X^{(1)},\ldots,X^{(k+1)}}$.

2.3 A Coefficient Based on a Different Prediction Rule

A rather striking property of $t_{Y\cdot \underline{X}}$ is that for $k = 2$, the reduction in prediction error equals just that corresponding to the more strongly associated $X^{(\nu)}$ of the two. Thus, $t_{Y\cdot \underline{X}}$ is a mathematically convenient but practically trivial measure when there are only two independent variables. The reason for this behavior lies with the prediction rule employed in formulating the coefficient. The rule of "predicting the majority ordering on Y with respect to the $\{X^{(\nu)}\}$" is a very simple one which leads to an easily interpretable coefficient. However, there is nothing unique about it, and more complex rules are necessary to produce a non-trivial measure when $k = 2$.

To formulate alternative coefficients of a nature similar to $t_{Y\cdot \underline{X}}$, one need only change the prediction rule. For example, suppose that a "proportional" prediction rule is utilized. That is, for a pair of observations in $A(\underline{\delta}) \cup A(-\underline{\delta})$, predict that

$$\underline{S}(i,j) = \underline{\delta} \text{ with probability } N(\underline{\delta})/(N(\underline{\delta}) + N(-\underline{\delta}))$$
$$\underline{S}(i,j) = -\underline{\delta} \text{ with probability } N(-\underline{\delta})/(N(\underline{\delta}) + N(-\underline{\delta})). \quad (2.14)$$

Then, the expected number of prediction errors for all such pairs in $A(\underline{\delta}) \cup A(-\underline{\delta})$ is $2N(\underline{\delta})N(-\underline{\delta})/(N(\underline{\delta}) + N(-\underline{\delta}))$, and considered over all such sets with fixed $\{X^{(\nu)}\}$ orderings, the proportional reduction in error using this rule is

$$t^{(2)}_{Y\cdot \underline{X}} = \frac{n(n-1)/2 - 2\sum_{D_k} N(\underline{\delta})N(-\underline{\delta})/(N(\underline{\delta}) + N(-\underline{\delta}))}{n(n-1)/2}. \quad (2.15)$$

It is easily verified that when $k = 1$, this coefficient reduces to the square of Kendall's tau (see Ploch (1974)). As the set of independent variables increases in size, $t^{(2)}_{Y\cdot \underline{X}}$ remains constant if the relative proportions used in the prediction remain unchanged. For example, $t^{(2)}_{Y\cdot X^{(1)},\ldots,X^{(k)}} = t^{(2)}_{Y\cdot X^{(1)},\ldots,X^{(k+1)}}$ if for all $\underline{\delta}$ in D_k

$$\frac{N(\delta_1,\ldots,\delta_k)}{N(-\delta_1,\ldots,-\delta_k)} = \frac{N(\delta_1,\ldots,\delta_k,1)}{N(-\delta_1,\ldots,-\delta_k,-1)} = \frac{N(\delta_1,\ldots,\delta_k,-1)}{N(-\delta_1,\ldots,-\delta_k,1)} \quad (2.17)$$

In particular, $t^{(2)}_{Y\cdot X^{(1)},X^{(2)}} > \max\left[t^{(2)}_{Y\cdot X^{(1)}}, t^{(2)}_{Y\cdot X^{(2)}}\right]$ unless

$$\frac{N(1)}{N(-1)} = \frac{N(1,1)}{N(-1,-1)} = \frac{N(1,-1)}{N(-1,1)}.$$

Thus, when $k = 2$, $t^{(2)}_{Y\cdot \underline{X}}$ is an especially useful measure of multiple rank association. For $k > 2$, $t^{(2)}_{Y\cdot \underline{X}}$ can be used as a supplementary measure to $t_{Y\cdot \underline{X}}$. However, it does not have quite as simple an interpretation as $t_{Y\cdot \underline{X}}$, and its value may seem somewhat artificial to the user, since it is naturally comparable to the squared rather that the unsquared Kendall's tau values. In essence, predictions based on this rule can result in no larger a reduction in error than those based on the rule previously described. Of course, $\sqrt{t^{(2)}_{Y\cdot \underline{X}}}$ could be used in comparison with the tau values, although then this coefficient lacks an interpretation. Also, the asymptotic moments and sampling distribution of $t^{(2)}_{Y\cdot \underline{X}}$ seem to be difficult to derive.

3. A MULTIPLE TAU COEFFICIENT FOR ORDINAL CATEGORICAL DATA

Tied pairs of observations would typically exist for most systems of variables in the social and behavioral sciences, where variables are commonly measured on ordinal categorical scales. If only a small proportion of pairs of observations are tied on at least one of the variables, one could continue to use $t_{Y\cdot \underline{X}}$ as defined in the previous section (tied pairs being omitted in the numerator). However, this results in a reduction in the potential magnitude of the measure, which becomes substantial as the proportion of tied pairs increases. For example, if the dependent variable is dichotomous with proportions .2 and .8 of observations in the two categories, the maximum possible value for $t_{Y\cdot \underline{X}}$ would be .32 (the proportion of pairs untied on Y) regardless of the distribution of ties among the independent variables.

3.1 Definition of $t'_{Y\cdot \underline{X}}$

To permit a maximum value of one and to ensure that the value does not decrease as independent variables are added to the system, one could base the coefficient on those pairs untied on Y and on at least one $X^{(\nu)}$, but standardize in the denominator according to the number of pairs untied on Y. That is, for $\delta_\nu = -1, 0$ or 1, $\nu = 1,\ldots,k$, let

$$A(\underline{\delta}) = \{(i,j):\ S[Y_j - Y_i] \neq 0 \text{ and } \underline{S}(i,j) = \underline{\delta}\} \quad (3.1)$$

and let $N(\underline{\delta})$ be the number of pairs of observations in $A(\underline{\delta})$. Let T_Y denote the number of pairs tied with respect to Y; i.e., if there are a_0 distinct values of Y with n_i observations tied at the i-th level, then $T_Y = \sum_{i=1}^{a_0} n_i(n_i - 1)/2$. Then, $\{A(\underline{\delta}),\ \delta_\nu = -1,0,1,\ \nu = 1,\ldots,k\}$ is a partition of the $n(n-1)/2 - T_Y$ pairs untied on Y, and $A(\underline{\delta}) \cup A(-\underline{\delta})$ is again the set of pairs with a fixed particular ordering on the $\{X^{(\nu)}\}$. Letting

$$D'_k = \{(\delta_1,\ldots,\delta_k):\ \delta_\nu = -1,\ 0, \text{ or } +1,\ \nu = 1,\ldots,k, \text{ at least one } \delta_\nu \neq 0\}, \quad (3.2)$$

we define

$$t'_{Y\cdot \underline{X}} = \frac{\frac{1}{2}\sum_{D'_k} |N(\underline{\delta}) - N(-\underline{\delta})|}{n(n-1)/2 - T_Y}. \quad (3.3)$$

Notice that $t'_{Y\cdot \underline{X}}$ may be rewritten as

$$t'_{Y\cdot \underline{X}} = \frac{\frac{1}{2}\left[\frac{n(n-1)}{2} - T_Y\right] - \frac{1}{2}\left[\sum_{D'_k} \min[N(\underline{\delta}), N(-\underline{\delta})] + N(0,\ldots,0)\right]}{\frac{1}{2}\left[\frac{n(n-1)}{2} - T_Y\right]} \quad (3.4)$$

That is, $t'_{Y\cdot \underline{X}}$ is the proportional reduction in error of predictions of the ordering on Y (for those pairs untied on Y) obtained by predicting majority ordering based on ordering of the $\{X^{(\nu)}\}$, relative to predicting randomly. Of course, when all $\delta_\nu = 0$, predictions in effect are also made randomly since the $\{X^{(\nu)}\}$ provide no predictive information, resulting in an expected number of errors of $N(0,\ldots,0)/2$.

Alternatively, $t'_{Y\cdot \underline{X}}$ may be written as

$$t'_{Y\cdot \underline{X}} = \sum_{D'_k} \lambda(\underline{\delta}) \frac{|N(\underline{\delta}) - N(-\underline{\delta})|}{N(\underline{\delta}) + N(-\underline{\delta})} = \frac{1}{2}\sum_{D'_k} (\lambda(\underline{\delta}) + \lambda(-\underline{\delta}))\,|t(\underline{\delta})|, \quad (3.5)$$

where $\lambda(\underline{\delta}) = \frac{N(\underline{\delta})}{n(n-1)/2 - T_Y}$ is the proportion of the pairs of observations untied on Y which are in $A(\underline{\delta})$. Thus, $t'_{Y\cdot \underline{X}}$ may be interpreted as a weighted average of the absolute values of the Kendall's taus within each set $A(\underline{\delta}) \cup A(-\underline{\delta})$ of orderings on the $\{X^{(\nu)}\}$, where a weight of $\frac{N(0,\ldots,0)}{n(n-1)/2 - T_Y}$ (the proportion of those pairs untied on Y which are tied on all $X^{(\nu)}$) is given to a value of $0 = N(0,\ldots,0) - N(0,\ldots,0)$. Here, $|t(\underline{\delta})|$ is the absolute value of Kendall's tau between Y and each of the $X^{(\nu)}$ such that $\delta_\nu \neq 0$, within the set of pairs $A(\underline{\delta}) \cup A(-\underline{\delta})$.

If $P(\underline{\delta})$ denotes the proportion of pairs of observations in $A(\underline{\delta})$ and p_i denotes the proportion of pairs tied on Y at the i-th of a_0 sets of ties on Y in some population of interest, then the population value of the coefficient $t'_{Y\cdot \underline{X}}$ is

$$\tau'_{Y\cdot \underline{X}} = \frac{\frac{1}{2}\sum_{D'_k} |P(\underline{\delta}) - P(-\underline{\delta})|}{1 - \sum_{i=1}^{a_0} p_i^2} = \frac{\sum_{D'_M} [P(\underline{\delta}) - P(-\underline{\delta})]}{1 - \sum_{i=1}^{a_0} p_i^2} \quad (3.6)$$

where $D'_M = \{\underline{\delta} \text{ in } D'_k : P(\underline{\delta}) > P(-\underline{\delta})\}$.

Alternative coefficients could again be formulated based on different prediction rules. For example, an extension of the proportional prediction rule discussed in the last section yields the multiple measure for ordinal categorical data,

$$t^{(2)'}_{Y\cdot\underline{X}} = \tag{3.7}$$

$$\frac{[n(n-1)/2 - T_Y] - 2 \sum_{D'_k \cup (0,\ldots,0)} N(\underline{\delta})N(-\underline{\delta})/(N(\underline{\delta}) + N(-\underline{\delta}))}{n(n-1)/2 - T_Y}.$$

3.2 Properties

Clearly, $t'_{Y\cdot\underline{X}}$ is invariant under strictly order preserving transformations on any of the variables. When there are no tied pairs with respect to any of the variables, $t'_{Y\cdot\underline{X}}$ reduces to the coefficient $t_{Y\cdot\underline{X}}$ discussed in Section 2. In the bivariate case, $t'_{Y\cdot X}$ reduces to Somers' d_{XY} (see Somers (1962)), a well-known asymmetric ordinal measure of association. When $k = 2$, $t'_{Y\cdot X^{(1)},X^{(2)}}$ is likely to be not much larger than $\max\{|t'_{Y\cdot X^{(1)}}|, |t'_{Y\cdot X^{(2)}}|\}$, but there is not necessarily equality here due to the additional contribution in the numerator of pairs tied on $X^{(1)}$ but not on $X^{(2)}$ and Y, or of pairs tied on $X^{(2)}$ but not on $X^{(1)}$ and Y. Again, though $t'_{Y\cdot\underline{X}}$ if of primary interest when $k \geq 3$, and a coefficient such as $t^{(2)'}_{Y\cdot\underline{X}}$ is likely to be of greater practical use when $k = 2$.

With the addition of $X^{(k+1)}$ to the set of independent variables,

$$t'_{Y\cdot X^{(1)},\ldots,X^{(k)}} \leq t'_{Y\cdot X^{(1)},\ldots,X^{(k+1)}},$$

since the denominator remains constant and the numerator can not decrease when the partition of pairs $\{A(\underline{\delta}) \cup A(-\underline{\delta})\}$ is refined.

For additional properties and a numerical example, see the related technical report on these measures (Agresti (1976)).

BIBLIOGRAPHY

[1] Agresti, A. (1976). Coefficients of Multiple Association for Ranked Data. Technical Report No. 112, University of Florida Department of Statistics.

[2] Ploch, D. R. (1974). Ordinal Measures of Association and the General Linear Model, in Hubert M. Blalock, ed., Measurement in the Social Sciences, Chicago: Aldine-Atherton, Ch. 12.

[3] Somers, R. H. (1962). A New Asymmetric Measure of Association for Ordinal Variables. American Sociological Review, 27, 799-811.