The Kendall Rank Correlation Coefficient

Hervé Abdi[1]

1 Overview

The Kendall (1955) rank correlation coefficient evaluates the degree of similarity between two sets of ranks given to a same set of objects. This coefficient depends upon the number of *inversions* of pairs of objects which would be needed to transform one rank order into the other. In order to do so, each rank order is represented by the set of all pairs of objects (*e.g.,* [a, b] and [b, a] are the two pairs representing the objects a and b), and a value of 1 or 0 is assigned to this pair when its order corresponds or does not correspond to the way these two objects were ordered. This coding schema provides a set of binary values which are then used to compute a Pearson correlation coefficient.

2 Notations and definition

Let $\mathscr{S}$ be a set of N objects

$$\mathscr{S} = \left\{ a, \ b, \ldots, x, y \right\} . \tag{1}$$

When the order of the elements of the set is taken into account we obtain an ordered set which can also be represented by the rank order given to the objects of the set. For example with the following set of $N = 4$ objects

$$\mathscr{S} = \{a,\ b,\ c,\ d\}\ , \tag{2}$$

the ordered set $\mathscr{O}_1 = [a, c, b, d]$ gives the ranks $\mathscr{R}_1 = [1, 3, 2, 4]$. An ordered set on N objects can be decomposed into $\frac{1}{2}N(N-1)$ ordered pairs. For example, $\mathscr{O}_1$ is composed of the following 6 ordered pairs

$$\mathscr{P}_1 = \{[a,c],\ [a,b],\ [a,d],\ [c,b],\ [c,d],\ [b,d]\}\ . \tag{3}$$

In order to compare two ordered sets (on the same set of objects), the approach of Kendall is to count the number of different pairs between these two ordered sets. This number gives a *distance* between sets (see entry on distance) called the *symmetric difference distance* (the symmetric difference is a set operation which associates to two sets the set of elements that belong to only one set). The symmetric difference distance between two sets of ordered pairs $\mathscr{P}_1$ and $\mathscr{P}_2$ is denoted $d_\Delta(\mathscr{P}_1, \mathscr{P}_2)$.

Kendall coefficient of correlation is obtained by normalizing the symmetric difference such that it will take values between -1 and $+1$ with -1 corresponding to the largest possible distance (obtained when one order is the exact reverse of the other order) and $+1$ corresponding to the smallest possible distance (equal to 0, obtained when both orders are identical). Taking into that the maximum number of pairs which can differ between two sets with

$$\frac{1}{2}N(N-1)$$

elements is equal to

$$N(N-1),$$

this gives the following formula for Kendall rank correlation coefficient:

$$\tau = \frac{\frac{1}{2}N(N-1) - d_\Delta(\mathscr{P}_1, \mathscr{P}_2)}{\frac{1}{2}N(N-1)} = 1 - \frac{2 \times [d_\Delta(\mathscr{P}_1, \mathscr{P}_2)]}{N(N-1)}\ . \tag{4}$$

How should the Kendall coefficient be interpreted? Because τ is based upon counting the number of different pairs between two ordered sets, its interpretation can be framed in a probabilistic context (see, *e.g.,* Hays, 1973). Specifically, for a pair of objects taken at random, τ can be interpreted as the difference between the probability for this objects to be in the same order [denoted $P(\text{same})$] and the probability of these objects being in a different order [denoted $P(\text{different})$]. Formally, we have:

$$\tau = P(\text{same}) - P(\text{different}). \tag{5}$$

3   An example

Suppose that two experts order four wines called $\{a, b, c, d\}$. The first expert gives the following order: $\mathcal{O}_1 = [a, c, b, d]$, which corresponds to the following ranks $\mathcal{R}_1 = [1, 3, 2, 4]$; and the second expert orders the wines as $\mathcal{O}_2 = [a, c, d, b]$ which corresponds to the following ranks $\mathcal{R}_2 = [1, 4, 2, 3]$. The order given by the first expert is composed of the following 6 ordered pairs

$$\mathcal{P}_1 = \{[a, c],\ [a, b],\ [a, d],\ [c, b],\ [c, d],\ [b, d]\} . \tag{6}$$

The order given by the second expert is composed of the following 6 ordered pairs

$$\mathcal{P}_2 = \{[a, c],\ [a, b],\ [a, d],\ [c, b],\ [c, d],\ [d, b]\} . \tag{7}$$

The set of pairs which are in only one set of ordered pairs is

$$\{[b, d]\ [d, b]\} , \tag{8}$$

which gives a value of $d_\Delta (\mathcal{P}_1, \mathcal{P}_2) = 2$. With this value of the symmetric difference distance we compute the value of the Kendall rank correlation coefficient between the order given by these two experts as:

$$\tau = 1 - \frac{2 \times [d_\Delta (\mathcal{P}_1, \mathcal{P}_2)]}{N(N-1)} = 1 - \frac{2 \times 2}{12} = 1 - \frac{1}{3} \approx .67 . \tag{9}$$

This large value of τ indicates that the two experts strongly agree on their evaluation of the wines (in fact their agree about everything but *one* pair). The obvious question, now, is to assess if such a large value could have been obtained by chance or can be considered as evidence for a real agreement between the experts. This question is addressed in the next section.

4   Significance test

The Kendall correlation coefficient depends only the order of the pairs, and it can always be computed assuming that one of the rank order serves as a reference point (*e.g.,* with $N = 4$ elements we assume arbitrarily that the first order is equal to 1234). Therefore, with two rank orders provided on N objects, there are $N!$ different possible outcomes (each corresponding to a given possible order) to consider for computing the sampling distribution of τ. As an illustration, Table 1, shows all the $N! = 4 \times 3 \times 2 = 24$ possible rank orders for a set of $N = 4$ objects along with its value of τ with the "canonical order" (*i.e.,* 1234). From this table, we can compute the probability p associated with each possible value of τ. For example, we find that the p-value associated with a one-tail test for a value of $\tau = \frac{2}{3}$ is equal to

$$p = P\left(\tau \geq \frac{2}{3}\right) = \frac{\text{Number of } \tau \geq \frac{2}{3}}{\text{Total number of } \tau} = \frac{4}{24} = .17 \ . \qquad (10)$$

We can also find from Table 1 that in order to reject the null hypothesis at the alpha level $\alpha = .05$, we need to have (with a one-tail test) perfect agreement or perfect disagreement (here $p = \frac{1}{24} = .0417$).

For our example of the wine experts, we found that τ was equal to .67, this value is smaller than the critical value of $+1$ (as given by Table 2) and therefore we cannot reject the null hypothesis, and we cannot conclude that the expert displayed a significant agreement in their ordering of the wines.

The computation of the sampling distribution is always theoretically possible because it is finite. But this requires computing

4

Table 1: The set of all possible rank orders for $N = 4$, along with their correlation with the "canonical" order 1234.

												Rank Orders												
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4
	2	2	3	3	4	4	1	1	3	3	4	4	1	1	2	2	4	4	1	1	2	2	3	3
	3	4	2	4	2	3	3	4	1	4	1	3	2	4	1	4	1	2	2	3	1	3	1	2
	4	3	4	2	3	2	4	3	4	1	3	1	4	2	4	1	2	1	3	2	3	1	2	1
τ	1	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0	0	$-\frac{1}{3}$	$\frac{1}{3}$	0	0	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{2}{3}$	0	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{2}{3}$	$-\frac{2}{3}$	-1

Table 2: Critical values of τ for $\alpha = .05$ and $\alpha = .01$.

	N = Size of Data Set						
	4	5	6	7	8	9	10
$\alpha = .05$	1	.8000	.7333	.6190	.5714	.5000	.4667
$\alpha = .01$	—	1	.8667	.8095	.7143	.6667	.6000

$N!$ coefficients of correlation, and therefore it becomes practically impossible to implement these computations for even moderately large values of N. This problem, however, is not as drastic as it seems because the sampling distribution of τ converges towards a normal distribution (the convergence is satisfactory for values of N larger than 10), with a mean of 0 and a variance equal to

$$\sigma_\tau^2 = \frac{2(2N+5)}{9N(N-1)} \ . \tag{11}$$

Therefore, for N larger than 10, a null hypothesis test can be performed by transforming τ into a Z value as:

$$Z_\tau = \frac{\tau}{\sigma_\tau} = \frac{\tau}{\sqrt{\dfrac{2(2N+5)}{9N(N-1)}}} \ . \tag{12}$$

This Z value is normally distributed with a mean of 0 and a standard deviation of 1.

For example, suppose that we have two experts rank ordering two sets of 11 wines, if the first expert gives the following rank order

$$\mathscr{R}_1 = [1,2,3,4,5,6,7,8,9,10,11]$$

and the second expert gives the following rank order

$$\mathscr{R}_2 = [1,3,4,5,7,8,2,9,10,6,11].$$

With these orders, we find that a value of $\tau = .6727$. When transformed into a Z value, we obtain

$$Z_\tau = \frac{.6727}{\sqrt{\dfrac{2(2 \times 10 + 5)}{9 \times 11 \times 10}}} = \frac{.6727}{\sqrt{\dfrac{54}{990}}} \approx 2.88 \ . \tag{13}$$

This value of $Z = 2.88$ is large enough to reject the null hypothesis at the $\alpha = .01$ level, and therefore we can conclude that the experts are showing a significant agreement between their evaluation of the set of wines.

4.1 Kendall and Pearson coefficients of correlation

Kendall coefficient of correlation can also be interpreted as a standard coefficient of correlation computed between two set of $N(N-1)$ binary values where each set represents all the possible pairs obtained from N objects and assigning a value of 1 when a pair is present in the order an 0 if not.

4.2 Extensions of the Kendall coefficient

Because Kendall rank order coefficient relies on a set distance, it can easily be generalized to other combinatoric structures such as weak orders, partial orders, or partitions. In all cases, the idea is similar: first compute the symmetric difference distances between the two set of the pairs representing the binary relation and then normalize this distance such that it will take values between -1 and $+1$ (see Degenne 1972, for an illustration and development of these ideas).

References

[1] Degenne, A. (1972) *Techniques ordinales en analyse des données.* Paris : Hachette.

[2] Hays, W.L. (1973). *Statistics.* New York: Holt Rinehart & Winston.

[3] Kendall, M.G., (1955). *Rank Correlation Methods.* New York: Hafner Publishing Co.

[4] Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGrawHill.