Kendall's Tau

Kendall's τ (tau) is a non-parametric measure of correlation between two ranked variables. It is similar to Spearman's ρ and Pearson's Product Moment Correlation Coefficient, or Pearson's r, in that is measures the relationship between two variables. Like Spearman's ρ and Pearson's r a negative correlation indicates that when X is increasing then Y is decreasing. Even though it is a similar to Spearman's ρ in that it is a non-parametric measure of relationship it differs in the interpretation of the correlation value. Spearman's ρ and Pearson's r magnitude are similar. However, Kendall's Tau represents a probability. In other words it is the difference between the probability that the observed data are in the same order versus the probability that the observed data are no in the same order. There are two variations of Kendall's Tau: tau-b and tau-c. They differ only in the way that they handle rank ties. This example show an example without any ties. As can be seen in Equation 1 there are many ways to show the equation. Equation 1 shows how Kendall's Tau is the probability of the difference of the concordant pairs and the discordant pairs. This is because the denominator is all possible combinations.

$$\tau = \frac{C-D}{C+D} = \frac{C-D}{\frac{n\cdot(n-1)}{2}} = \frac{C-D}{\binom{n}{2}} = \frac{C-D}{\frac{n!}{2!(n-2)!}} \tag{1}$$

Where:

C = Concordant Pairs

D = Discordant Pairs

A concordant pair is when the rank of the second variable is greater than the rank of the former variable. A discordant pair is when the rank is equal to or less than the rank of the first variable.

Table 1: Kendall's Tau Rank Correlation Without Ties

Rank 1	Rank 2
1	1
2	3
3	6
4	2
5	7
6	4
7	5

Calculating Kendall's Tau manually can be very tedious without a computer and is rarely done without a computer. Large dataset make it almost impossible to do by manually by hand.

Table 2: Counting Concordant and Discordant Values for Kendall's Tau

R2							
1							
2	C						
3	C	C					
4	C	D	D				
5	C	C	C	C			
6	C	C	C	D	D		
7	C	C	C	C	D	D	
	1	2	3	4	5	6	7

To calculate the C and D simply count up the total number of C's and D's. These values are then used in Equation 2.

$$\tau = \frac{15-6}{15+6} = \frac{7}{21} = .42857 \tag{2}$$

The process to calculate Kendall's τ is quite simple using the R Environment. The *cor* command will provide Pearson's r, Spearman's ρ, and Kendall's τ. Significance testing can be computed using the *Kendall* library.

```
x <- c(1,2,3,4,5,6,7);
y <- c(1,3,6,2,7,4,5);
cor(x,y, method="kendall");

##For significance testing
library("Kendall");
summary(Kendall(x,y));
```