International J.Math. Combin. Vol.1(2012), 74-77

Combinatorial Aspects of a Measure of Rank Correlation Due to Kendall and its Relation to Complete Signed Digraphs

P.Siva Kota Reddy, Kavita S.Permi and K. R.Rajanna

Department of Mathematics, Acharya Institute of Technology, Chikkabanavara, Bangalore-560 090, India

E-mail: pskreddy@acharya.ac.in

Abstract: In this paper, we shall present an account of certain combinatorial aspects of a measure of rank correlation due to Kendall (1938) and point out its relation to the analysis of patterns of preference and indifference which in recent years, have been matters of intense discussion among the social psychologists because of their fundamental role in dealing with certain vital issues of social decision theory.

Key Words: Smarandachely k-signed graph, Smarandachely k-marked graph, signed digraph, rank, Kendalls τ.

AMS(2010): 05C22

§1. Introduction

For standard terminology and notion in digraph theory, we refer the reader to the classic textbooks of Bondy and Murty [1]and Harary et al. [3]; the non-standard will be given in this paper as and when required.

A *Smarandachely k-signed graph* (*Smarandachely k-marked graph*) is an ordered pair $S = (G, \sigma)$ ($S = (G, \mu)$) where $G = (V, E)$ is a graph called *underlying graph of* S and $\sigma : E \rightarrow (\overline{e}_1, \overline{e}_2, ..., \overline{e}_k)$ ($\mu : V \rightarrow (\overline{e}_1, \overline{e}_2, ..., \overline{e}_k)$) is a function, where each $\overline{e}_i \in \{+, -\}$. Particularly, a Smarandachely 2-signed graph or Smarandachely 2-marked graph is called a *signed graph* or a *marked graph*. A signed digraph $S = (D, \sigma)$ is *balanced*, if every semicycle of S is positive (See [3]). Equivalently, a signed digraph is balanced if every semicycle has an even number of negative arcs. The following characterization of balanced signed digraphs is obtained in [6].

Proposition 1.1(E. Sampathkumar et al. [6]) *A signed digraph $S = (D, \sigma)$ is balanced if, and only if, there exist a marking μ of its vertices such that each arc $\overrightarrow{uv}$ in S satisfies $\sigma(\overrightarrow{uv}) = \mu(u)\mu(v)$.*

In [6], the authors also introduced the switching and cycle isomorphism for signed digraphs. The *rank* means the position of an item or datum in relation to others which have been arranged according to some specific criterion, when used as verb, it means the act of assigning a rank to

[1]Received September 13, 2011. Accepted March 5, 2012.

each term or datum according to the specified criterion (See, Wolman [7]). Ranking is then the arrangement of a series of values, scores or individuals in the order of their magnitude which may be decreasing or increasing. The problem of ranking individuals according to the extent of a prescribed attribute possessed or exhibited by them is of primary importance in the process of interpretation of statistical data for decision making in a wide variety of situations arising in experimental behavioral sciences. Moreover, when the individuals are ranked separately with regard to two different attributes, the two rankings may not be the same in general and hence the problem of effectively comparing the two rankings arises. A natural approach to such comparison is to quantify the content of correlation between the two rankings in some way that would reflect the individuals standing in each of the two rankings. Such a numerical value used to represent the content of correlation that may exist between any two given rankings is called a *measure of rank correlation*.

Several types of rank correlation measures exist in literature (See, Guilford and Fruchter [2], Kendall [4]). Of particular interest to us in this paper is Kendall τ ('tau') measure which resets on no regression analytic assumptions (See, Kendall [4]). This measure has numerous applications, including the testing of hypotheses, but bears no direct relation to the traditional family of productmoment correlations (See, Guilford and Fruchtter [2]).

§2. Complete Signed Digraphs and of Measure of Rank Correlation

Given two rankings A and B of n individuals $v_1, v_2, ..., v_n$, Kendall [4] has defined a new measure τ of rank correlation between A and B. We give here a method of construction of a complete signed digraph $S = (D, \sigma)$ with n vertices, from which we can easily find Kendall s τ by the formula $\tau = \frac{P-N}{P+N}$, where P and N respectively denote the number of positive arcs and number of negative arcs in S. A complete signed digraph is a complete digraph in which every arc is assigned either $+$ or $-$.

In [5], Sampathkumar and Nanjundaswamy obtained Kendall s τ for complete signed graphs. By the motivation of Kendalls τ for complete signed graphs, here we make an attempt to obtain the same for complete signed digraphs. Let $V = v_1, v_2, ..., v_n$ and a ranking of V is a bijective map $A : V \to \{1, 2, ..., n\}$.

Let A and B be two rankings of $V = \{v_1, v_2, ..., v_n\}$. Construct a signed digraph S on the complete digraph $\overrightarrow{K_n}$ whose vertices are labeled $v_1, v_2, ..., v_n$ as follows: consider A as the objective ranking and change the label of $\overrightarrow{K_n}$ according to the rule: $v_i \to V_{A(v_i)}$, for all $i \in \{1, 2, ..., n\}$. Observe that $B(V_j) = B(v_j)$, whenever $A(v_i) = j$. Now, label the arcs of $\overrightarrow{K_n}$ with respect to the new labeling recursively as below. For each $i = 1, 2, ..., n-1$,

$$\sigma(\overrightarrow{V_iV_j}) = \begin{cases} +, & \text{if } B(V_i) < B(V_j) \\ -, & \text{otherwise} \end{cases}$$

for each j, $j = i+1, i+2, ..., n$. From the above, we can easily observe that

$$\sigma(\overrightarrow{V_jV_i}) = \begin{cases} +, & \text{if } B(V_j) < B(V_i) \\ -, & \text{otherwise} \end{cases}$$

Hence, let P and N denote the number of positive and negative arcs respectively in the signed digraph S so obtained. Then the Kendalls measure τ of rank correlation is given by $\tau = \dfrac{P-N}{P+N}$. We now illustrate the above method of finding τ by an example.

Example 2.1 Let A and B be the two rankings of four individuals v_1, v_2, v_3 and v_4 given as follows:

Individuals	v_1	v_2	v_3	v_4
Ranking A :	3	1	4	2
Ranking B :	1	2	3	4

Draw the complete signed digraph S on v_1, v_2, v_3 and v_4. Consider A as the objective ranking. Relabel the vertices of S as follows: $v_1 \to V_3$, $v_2 \to V_1$, $v_3 \to V_4$ and $v_4 \to V_2$.

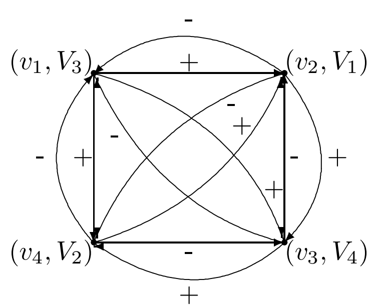

Fig.1

In Fig.1, $B(V_3) = B(v_1) = 1$, $B(V_1) = B(v_2) = 2$, $B(V_4) = B(v_3) = 3$ and $B(V_2) = B(v_4) = 4$. Here, we sign the arcs of D as mentioned above. For example, since $B(V_1) > B(V_3)$ ($B(V_3) < B(V_1)$) the arc $e = \overrightarrow{V_1V_2}$ ($e = \overrightarrow{V_2V_1}$) is labeled as negative (positive).

The signed digraph obtained in the above figure has $P = 6$ and $N = 6$. Hence, $\tau = 0$.

We can easily deduce the following properties of τ from the formula $\tau = \dfrac{P-N}{P+N}$:

1. $|\tau| \leq 1$

2. There is a perfect positive correlation between the rankings of A and $B \Leftrightarrow$ all arcs in S are positive, i.e, $N = 0 \Leftrightarrow \tau = 1$.

3. The rankings A and B are exactly inverted $\Leftrightarrow$ all arcs are negative, i.e, $P = 0 \Leftrightarrow \tau = -1$.

4. If the arc $\overrightarrow{V_1V_n}$ in the signed digraph is positive (negative) then it follows by the construction of that $P \geq N$ ($P \leq N$) and hence $\tau \geq 0$ ($\tau \leq 0$).

5. Since each arc of the signed digraph S is labeled either positive or negative, the probability that an arc is positive (negative) is $\frac{1}{2}$, and hence the number of positive (negative) arcs is a binomial variate. Since the probability that an arc is positive is exactly equal to the probability that is negative, the limiting case of distribution of number of individuals, is

increased indefinitely, is the normal distribution. Hence, the distribution of τ tends to normality for large n.

The original Kendalls method of finding τ goes as follows: Rearrange A as an objective order A' and write below it corresponding ranks in B to obtain the relative permutation $B' = (b'_1, b'_2, ..., b'_n)$. For each j, $j = 1, 2, ..., n-1$, consider the order of allot $n-j$ ordered pairs (b'_j, b'_k), $k = j+1, ..., n$ and allot score $+1$ to (b'_j, b'_k) if $b'_j < b'_k$ (then we say that (b'_j, b'_k) is in the correct order) and allot a score -1 to (b'_j, b'_k) if $b'_j > b'_k$ (then we say that (b'_j, b'_k) is in the wrong order). The sum of all these scores is called the *actual score*. The actual score of the objective order $A' = (1, 2, ..., n)$ is thus the combinatory function n_{C_2} and obviously must be the maximum possible score for any ranking of n items. Kendalls definition of τ is then,

$$\tau = \frac{\text{actualscore}}{\text{maximumscore}}.$$

Applying this procedure to the rankings A and B of above Example, we find and then computing τ according to $\tau = \frac{\text{actualscore}}{\text{maximumscore}}$ it turns out that $\tau = 0$, the value which also came out by computing according to $\tau = \frac{P-N}{P+N}$. That in general $\tau = \frac{P-N}{P+N}$ and $\tau = \frac{\text{actualscore}}{\text{maximumscore}}$ are equivalent.

Individuals	v_1	v_2	v_3	v_4
Ranking A :	1	2	3	4
Ranking B :	2	4	1	3

Acknowledgement

The authors would like to thank referee for his constructive suggestions and remarks which improved the quality of presentation. Also, the authors is grateful to Sri. B. Premnath Reddy, Chairman, Acharya Institutes, for his constant support and encouragement for research and development.

References

[1] J.A.Bondy and U.S.R.Murty, *Graph Theory with Applications*, MacMillan, London, 1976.

[2] J.P.Guilford and B.Fruchter, *Fundamentals Statistics in Psychology and Education*, Mc-Graw Hill-Kogakusha Ltd., Tokyo, 1973.

[3] F.Harary, R. Norman and D. Cartwright, *Structural Models: An Introduction to the Theory of Directed Graphs*, J. Wiley, New York, 1965.

[4] M. G.Kendall, A new measure of rank correlation, *Biometrica*, Vol.XXX (1938), 81-93.

[5] E.Sampthkumar and L.Nanjundaswamy, Complete signed graphs of mesure of rank correlation, *The Karnatak University Journal: Science*, Vol.XVIII (1973), 308-311.

[6] E.Sampathkumar, M. S.Subramanya and P.Siva Kota Reddy, Characterization of Line Sidigraphs, *Southeast Asian Bull. Math.*, 35(2) (2011), 297-304.

[7] B.Wolmen, *Dictionary of behavioural sciences*, Van-Nostrend Reinhold Company, New York, 1973.