2011 International Conference on Advancements in Information Technology
With workshop of ICBMG 2011
IPCSIT vol.20 (2011) © (2011) IACSIT Press, Singapore

Quantitative Measurement of Scores by Ranks

M. Gayatri [1+] and MHM Krishna Prasad[2]

[1] Dept. of CSE, UCEV – JNTUK, Vizianagaram, A.P., India

[2] Associate Professor & Head, Dept. of IT, UCEV – JNTUK, Vizianagaram, A.P., India

Abstract. Ranking, which to have a position higher or lower than others and consists of listing of items in a group, such as schools or sports teams, according to a system of rating or a record of performance. It is performed on the basis of some scores available for each item. Comparing two different rankings has been studied in various fields. In each case, a measure has been provided that takes into account how much the positions of each item differ in the two ordered lists. The existing practice for comparing scoring functions is to compare the induced rankings by one of a large number of rank comparison methods available in the literature. In our paper we demonstrate that our approach outperforms the previously proposed automatic ranking method. For comparing the similarity of rankings of two search engines, the Spearman correlation coefficient is computed. When comparing more than two sets Kendall's W is used. We suggest that it may be better to compare the underlying scores themselves. To this end, a generalized Kendall distance is defined .which takes into consideration not only the final ordering that the two schemes produce but also at the spacing between pairs of scores. Experimental results clearly show the advantages score comparison has over rank comparison.

Keywords: Kendall tau Distance, spearman's correlation coefficient, rank,

1. Introduction

The Kendall rank correlation coefficient commonly referred to as Kendall's tau (T) coefficient is a statistic used to measure the association between two measured quantities. Specifically, it is a measure of rank correlation that is the similarity of the data when ranked by each of the quantities. The Kendall tau [2] distance is a metric that counts the number of pair wise disagreements between two lists. The larger the distance, the more dissimilar the two lists are. The Kendall tau distance. $K(\tau 1,\tau 2)$ will be equal to 0 if the two lists are identical and $n(n - 1) / 2$ (where n is the list size) if one list is the reverse of the other. Often Kendall tau distance is normalized by dividing by $n(n - 1) / 2$ so a value of 1 indicates maximum disagreement. The normalized Kendall tau distance therefore lies in the interval [0,1]. T= (number of concordant pairs)-(no. of Discordant pairs)/1/2n(n-1).

Example Suppose we rank a group of five people by height and by weight:

Table 1: Rank a group of five people by height and by weight

Person	A	B	C	D	E
Rank by Height	1	2	3	4	5
Rank by Weight	3	4	1	2	5

Here person A is tallest and third-heaviest, and so on. In order to calculate the Kendall tau distance, pair each person with every other person and count the number of times the values in list 1 are in the opposite order of the values in list 2.

Table 2: List of values for height and weight

Pair	Height	Weight	Count
(A,B)	1 < 2	3 < 4	

+ Corresponding author. Tel.: +91 9293559948.
E-mail address: gayatri1201@gmail.com.

(A,C)	1 < 3	3 > 1	**X**
(A,D)	1 < 4	3 > 2	**X**
(A,E)	1 < 5	3 < 5	
(B,C)	2 < 3	4 > 1	**X**
(B,D)	2 < 4	4 > 2	**X**
(B,E)	2 < 5	4 < 5	
(C,D)	3 < 4	1 < 2	
(C,E)	3 < 5	1 < 5	
(D,E)	4 < 5	2 < 5	

Since there are 4 pairs whose values are in opposite order, the Kendall tau distance is 4.

4/5(5-1)/2=0.4

A value of 0.4 indicates a somewhat low agreement in the rankings.

2. Kendall tau Rank Correlation Coefficient

In statistics, the Kendall rank correlation coefficient commonly referred to as Kendall's tau (τ) coefficient, is a statistic used to measure the association between two measured quantities. A tau test is a non-parametric hypothesis test which uses the coefficient to test for statistical dependence. Specifically, it is a measure of rank correlation: that is, the similarity of the orderings of the data when ranked by each of the quantities. The Kendall tau distance is a metric that counts the number of pair wise disagreements between two lists. The larger the distance, the more dissimilar the two lists are. Kendall tau distance is also called bubble-sort distance since it is equivalent to the number of swaps that the bubble sort algorithm would make to place one list in the same order as the other list. The Kendall tau distance was created by Maurice Kendall. The Kendall tau distance between two lists τ1 and τ2 is

$$K(\tau_1,\tau_2) = |(i,j) : i < j, (\tau_1(i) < \tau_1(j) \wedge \tau_2(i) > \tau_2(j)) \vee (\tau_1(i) > \tau_1(j) \wedge \tau_2(i) < \tau_2(j))|.$$

K(τ1,τ2)=|(I,J):I<j,()

K(τ1,τ2) will be equal to 0 if the two lists are identical and n(n − 1) / 2 (where n is the list size) if one list is the reverse of the other. Often Kendall tau distance is normalized by dividing by n(n − 1) / 2so a value of 1 indicates maximum disagreement. The normalized Kendall tau distance therefore lies in the interval [0,1].

Kendall tau distance may also be defined as

$$K(\tau_1, \tau_2) = \sum_{\{i,j\}\in P} \bar{K}_{i,j}(\tau_1, \tau_2)$$

Where.

- P is the set of unordered pairs of distinct elements in τ1 and τ2
- $\bar{K}_{i,j}(\tau_1, \tau_2)$ = 0 if *i* and *j* are in the same order in τ1 and τ2
- $\bar{K}_{i,j}(\tau_1, \tau_2)$ = 1 if *i* and *j* are in the opposite order in τ1 and τ2.

Kendall tau distance can also be defined as the total number of discordant pairs.

Let *(x1, y1), (x2, y2),... (xn, yn)* be a set of joint observations from two random =variables X and Y respectively, such that all the values of *(xi)* and *(yi)* are unique. Any pair of observations *(xi, yi)* and *(xj, yj)* are said to be concordant if the ranks for both elements agree: that is, if both *xi > xj* and *yi > yj* or if both *xi < xj* and *yi < yj*. They are said to be discordant, if *xi > xj* and *yi < yj* or if *xi < xj* and*yi > yj*. If *xi = xj* or *yi = yj*, the pair is neither concordant nor discordant.

The Kendall τ coefficient is defined as:

$$\tau = \frac{(\text{number of concordant pairs}) - (\text{number of discordant pairs})}{\frac{1}{2}n(n-1)}.$$

The properties of the kendall tau distance as, The denominator is the total number of pairs, so the coefficient must be in the range −1 ≤ τ ≤ 1.If the agreement between the two rankings is perfect (i.e., the two rankings are the same) the coefficient has value 1.If the disagreement between the two rankings is perfect (i.e., one ranking is the reverse of the other) the coefficient has value −1.If X and Y are independent, then we would expect the coefficient to be approximately zero.

3. Algorithms

The direct computation of the numerator nc − nd, involves two nested iterations, as characterized by the following pseudo-code:

```
number := 0;
for i:=1..N do
for j:=1..(i-1) do
number := number + sign(x_i-x_j)*sign(y_i-y_j);
return number;
```

Although quick to implement, this algorithm is O(n2) in complexity and becomes very slow on large samples. A more sophisticated algorithm built upon the Merge Sort algorithm can be used to compute the numerator in time.

Begin by ordering your data points sorting by the first quantity, x, and secondarily (among ties in x) by the second quantity, y. With this initial ordering, y is not sorted, and the core of the algorithm consists of computing how many steps a Bubble Sort would take to sort this initial y. An enhanced Merge Sort algorithm, with O(nlogn) complexity, can be applied to compute the number of swaps, *S(y)*, that would be required by a Bubble Sort to sort yi. Then the numerator for τ is computed as: $nc - nd = n0 - n1 - n2 + n3 - 2S(y)$, where n3 is computed like n1 and n2, but with respect to the joint ties in x and y.

A Merge Sort partitions the data to be sorted, y into two roughly equal halves, y_{left} and y_{right}, then sorts each half recursive, and then merges the two sorted halves into a fully sorted vector. The number of Bubble Sort swaps is equal to: $S(y) = S(y_{left}) + S(y_{right}) + M(Y_{left}, Y_{right})$, where Y_{left} and Y_{right} are the sorted versions of y_{left} and y_{right}, and characterizes the Bubble Sort swap equivalent for a merge operation. is computed as depicted in the following pseudo-code:

```
M(L,R)
{
n := |L| + |R|;
i := 1;
j := 1;
nSwaps := 0;
while (i+j <= n) do
{
if i>|L| or R[j] < L[i] then
nSwaps := nSwaps + |L|- (i-1);
j := j + 1;
else
i := i + 1;
end if
);
return nSwaps;
}
```

A side effect of the above steps is that you end up with both a sorted version of x and a sorted version of y. With these, the factors *ti* and *uj* used to compute τB are easily obtained in a single linear-time pass through the sorted arrays.

3.1. Spearman's Rank Correlation Coefficient

It is a non-parametric measure of statistical dependence between two variables. If there are no repeated data values, a perfect spearman correlation of +1 or -1 occurs. In statistics, Spearman's rank correlation coefficient or Spearman's rho, named after Charles Spearman and often denoted by the Greek letter ρ (rho) or as *rs*, is a non-parametric measure of statistical dependence between two variables. It assesses how well the relationship between two variables can be described using a monotonic function. If there are no repeated data values, a perfect Spearman correlation of +1 or −1 occurs when each of the variables is a perfect monotone function of the other.

The Spearman correlation coefficient is defined as the Pearson correlation coefficient between the ranked variables.[1] The n raw scores X_i, Y_i are converted to ranks x_i, y_i, and ρ is computed from these:

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}.$$

We can determine the significance of Spearman correlation coefficient. One approach to testing whether an observed value of ρ is significantly different from zero (r will always maintain $1 \geq r \geq -1$) is to calculate the probability that it would be greater than or equal to the observed r, given the null hypothesis, by using a permutation test. An advantage of this approach is that it automatically takes into account the number of tied data values there are in the sample, and the way they are treated in computing the rank correlation.

Another approach parallels the use of the Fisher transformation in the case of the Pearson product-moment correlation coefficient. That is, confidence intervals and hypothesis tests relating to the population value ρ can be carried out using the Fisher transformation:

$$F(r) = \frac{1}{2} \ln \frac{1+r}{1-r} = \operatorname{arctanh}(r).$$

If $F(r)$ is the Fisher transformation of r, the sample Spearman rank correlation coefficient, and n is the sample size, then

$$z = \sqrt{\frac{n-3}{1.06}} F(r)$$

is a z-score for r which approximately follows a standard normal distribution under the null hypothesis of statistical independence ($\rho = 0$).[5, 6]

One can also test for significance using

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

which is distributed approximately as Student's t distribution with $n - 2$ degrees of freedom under the null hypothesis[7]. A justification for this result relies on a permutation argument.[8] A generalization of the Spearman coefficient is useful in the situation where there are three or more conditions, a number of subjects are all observed in each of them, and it is predicted that the observations will have a particular order. For example, a number of subjects might each be given three trials at the same task, and it is predicted that performance will improve from trial to trial. A test of the significance of the trend between conditions in this situation was developed by E. B. Page [9] and is usually referred to as Page's trend test for ordered alternatives.

4. Results and Analysis

One may compare two score vectors directly using a measure like Pearson's correlation coefficient. However, the interpretation of the coefficient in terms of the resultant ranking is lost. Also, the correlation coefficient is not a metric and hence cannot be interpreted directly as a distance between two scoring systems. The correlation coefficient may be transformed into a metric, but it still does not reflect the rank specific differences between two scorings and is not always useful for comparing rank-inducing scoring functions.

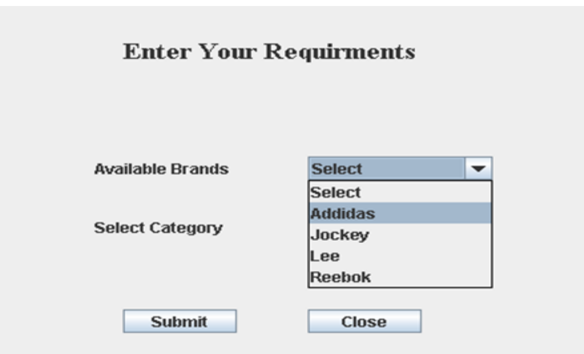

Fig 1: Entering the requirements

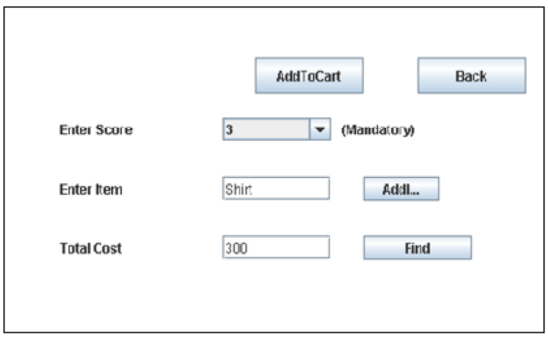

Fig 2: Add to kart to get ranks

A top k list [3] is the set of items with the largest scores. Top k lists differ from full lists because two lists need not have the same set of items. We first study the approach of Fagin et al. [10] for generalizing the definition of discordance to the case of top k lists. We reproduce the text from [10] and, simultaneously, make a note of how the same extension for computing the degree of discordance would differ in each case. Let *s1* and *s2* be two top k lists. Ranking module is used for ranking a set of items is a fairly frequent task and involves pair wise comparison of the given items. This comparison may be performed by inquiring an oracle for each pair of items, in which case, the ranking procedure is known as comparison-based ranking. Ranking web pages has attracted the attention of several researchers, mainly due to the challenges it poses in terms of scalability and the imprecise and subjective nature of the task.

5. Conclusion and Future Work

This paper describes generalizing measures of discordance for the case when the underlying scores are known. A metric has been provided to compare score vectors directly. This metric turns out to be the Kendall distance when a parameter denoting the ratio of fusing proportions, is large. Experiments of various kinds demonstrate the wide range of theory and applications of the metric introduced in the present work. There is a tremendous scope for future work, including studying the cases where T is assumed to arise from specific distributions, obtaining the properties such as maximum and minimum. For particular values and speeding up the computation of the proposed metric.

6. References

[1] F. Crestani. Combination of Similarity Measures for Effective Spoken Document Retrieval. Information Science, vol. 29, no. 2, pp. 87-96, 2003.

[2] W.R. Knight, "A Computer Method for Calculating Kendall's Tau with Ungrouped Data," J. Am. Statistical Assoc., vol. 61, no. 314,pp. 436-439, 1966.

[3] R. Fagin, R. Kumar, and D. Sivakumar, "Comparing Top k Lists," Siam J. Discrete Math., vol. 17, no. 1, pp. 134-160, 2003.

[4] M.E. Renda and U. Straccia. Web Metasearch: Rank versus Score Based Rank Aggregation Methods. In: Proc. 18th Ann. ACM Symp. Applied Computing (SAC'03). 2003, pp. 841-846.

[5] W.J. Conover. Practical Nonparametric Statistics. 3 ed, John Wiley & Sons, 1999.

[6] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar. Rank Aggregation Methods for the Web. In: Proc. 10th Int'l World Wide Web Conf. (WWW '01), pp. 613-622, 2001.

[7] Press, Vettering, Teukolsky, and Flannery. Numerical Recipes in C: The Art of Scientific Computing. 2nd Edition, page 640, 1992.

[8] Kendall, M.G, Stuart. The Advanced Theory of Statistics, Volume 2: Inference and Relationship, Griffin. ISBN 0852642156, 1973.

[9] Page, E. B., Ordered hypotheses for multiple treatments: A significance test for linear ranks. Journal of the American Statistical Association 58: 216–230. doi:10.2307/2282965, 1963.

[10] Bhamidipati, N.L. Pal, S.K. Comparing Scores Intended for Rankin, Knowledge and Data Engineering, IEEE Transactions on, Jan. 2009 Volume: 21 Issue: 1, 21 – 34, ISSN: 1041-4347.