Derivation the Multiple Correlation Coefficient of Ranks by the Ranks Determinant

Abdulrahman Zeyarah
Al-Furat Al-Awsat Technical University

Sabah Mohammed Al-Mutoki
Al-Furat Al-Awsat Technical University

Derivation the Multiple Correlation Coefficient of Ranks by the Ranks Determinant

Abdulrahman Kadum Abbas Zeyarah, Department of Electricity Techniques, Al-Furat Al-Awsat Technical University, Technical Institute / Samawa- Iraq. E-mail: abidlrahman.zeyareh@yahoo.com

Evan Abdulkareem Huzam, Thi-Qar University, Faculty of Education for Pure Science. E-mail: evan_krm@yahoo.com

Sabah Mohammed Mlket Almutoki, Department of Electricity Techniques, Al-Furat Al-Awsat Technical University, Technical Institute / Samawa- Iraq. E-mail: asbah_sh2003@yahoo.com

Abstract--- The rank correlation coefficient between two observed values is computed by, “Spearman rank correlation coefficient”. In this paper, it has been inferred formula of the Multiple Correlation Coefficient of ranks, when the ranks of the observed values with more than two variables, make up determinant, called the ranks determinant. The value of the ranks determinant indicates the amount of difference between the column ranks, and these differences are statistically significant.

Key Statements--- Multiple Correlation Coefficient of Ranks, Ranks Determinant, Generator Variable, Function of the Generator Variable

I. Introduction

The rank correlation coefficient between two observed values is computed by, Spearman rank correlation coefficient[1], however, this coefficient is not for more than two variables, and it is well known, that the “Pearson multiple correlation coefficient” describes the correlation between observed values, not between their ranks. In this paper, will carrying out the finding of the *Multiple Correlation Coefficient of Ranks*, which is possible to found it when the observed values compose squares matrix. For instead, in questionnaires when the number of questions equal to the number of the options.

(1.1) The Algorithm: in order to obtain a database related to the subject of the paper, an algorithm must be followed by the extraction of the data, and in order to derive the possible relationships and information. Therefore, the search will depend on the following algorithm.

(1.1.1). Designation of the r setas a positive numbers, as the ranks of the observed values.

(1.1.2).Finding the permutation of r, and listed them as the columns.

(1.1.3). Limitation all the possible of the determinants by taking $r$columns from the set of columns, which are listed in step.(1.1.2).

(1.1.4). Evaluating the value of the determinants, which are listed in step. (1.1.3).

(1.1.5). Specify the different values of the determinants, as all repeated values are a single value. In this step, we’ll get regular values in arithmetic sequence; it represents the most important database information.

(1.2)The Generator Variable: we called the number r is the *Generator Variable*, it is responsible for generating many relationships and all the statistics related to the subject of this research, will be evident in the context of the paper. In addition, from finite number$n = 1,2, \dots, r$.a matrix can be created, it size is $r \times r$, and each column (row) contains r elements.

Since the generator variable can be taken any finite number, the size of Ranks Determinant is finite too.(Sometimes will be referred to Ranks Determinant by the short symbol $R_S D$ in singular).

For sufficient information, starting let us limit our attention to three ranks of the observed values as the following:

i) Three ranks: $r = 1, 2, 3$.For 3×3 size of the $R_S D$.
ii) Four ranks: $r = 1, 2, 3, 4$.For 4×4 size of the $R_S D$.
iii) Five ranks: $r = 1, 2, 3,4 ,5$.For 5×5 size of the $R_S D$.

Received: 16 July 2019/Accepted: 05 August 2019

And so on.

For example, three options, four options, and five options in questionnaires, and so on.

II. Ranks Determinant

(2.1)Definition: The *Ranks Determinant* is the determinant of the square matrix, which its elements a_{ij} are the ranks of the observed values, with size *r*, such that$r \geq 3$, and finite positive number.

Usually, the ranks determinant are natural numbers, or positive rational numbers, which are refers to rank of each observed value.

(2.2)The Different Values and Maximum Values of the Ranks Determinant: in during reviewing of the different values of the Ranks Determinants, for size $r = 3,4, and\ 5$,it found the following arithmetic sequences.

i) If the size of the $R_S D$ is $r = 3$; the ranks of the observed values are $r = 1, 2, 3$; then the listing the positive values of the $R_S D$ given by the following arithmetic sequence: $0, 6, 12, 18$. Or in term the generator variable *r* is

$$0, 2r, 4r, 6r. \dots \dots (1a)$$

ii) If the size of the $R_S D$ is $r = 4$; the ranks of the observed values are $r = 1,2,3,4$; then the listing the positive values of the ranks determinant given by the following sequence: $0, 10, 20, 30, \dots, 160$. Or in term the generator variable r is

$$0, 2(r + 1), 4(r + 1), 6(r + 1), \dots, 32(r + 1) \ \dots \dots (1b)$$

iii) If the size of the $R_S D$ is $r = 5$; the ranks of the observed values are $r = 1, 2, 3, 4, 5$. Then the listing the positive values of the $R_S D$ given by the following sequence:

$0, 15, 30, 45, \dots, 1875$. Or in term the generator variable *r* is

$$0, 3r, 6r, 9r, 12r, 15r, \dots, 125r, 375r. \dots \dots (1c)$$

The (1a), (1b), and (1c) are arithmetic sequences, it is easy to note that the last term in any arithmetic sequence above is the maximum positive value of the $R_S D$, and they are denoted in term r by symbol: $Max|Det(A)|_r$, such that $|Det(A)|$ is absolute value of the ranks determinant *A*. See the table (2.1) below:

Table 2.1: The absolute maximum Values of Ranks Determinants

Size (*r*)	3	4	5
$Max|Det(A)|$	18	160	1875

(2.3) The Number of the Different Values of the Ranks Determinant: the permutation of $r$elements, which represent the ranks of the observed values equal to $r!$.Suppose $r! = n$, such that $r!$the number of the columns of ranks, then the number of $R_S D$, which are produced from permutation *n* columns choice r columns without repeat is given by the following formula:$n\ p\ r = n(n-1)(n-2)(n-3) \dots (n - r1) = \frac{n!}{(n-1)!}$[2],this number of Ranks determinants are contains some ranks determinants, which their values are equal, naturally regarding them only one $R_S D$ in term of value.

In reviewing the values of the ranks determinants, it found that the *number of the different values of the Ranks Determinant*, in term *r* equal to:

$$N = 2r^{r-2} + 1 \dots \dots (2a)$$

Number one in formula.(2a) refers to the zero value of $R_S D$, and the reminder numbers are divided into two equality sets numbers; positive and negative values of $R_S D$.And every value has the inverse, for instance; when $r = 3$: one value equal to zero, 3 positive values, and 3 negative values. The ranks determinants, which are have zero value may they are *Progressive Elements Determinant*, with common difference equal to 1, since its size $r \geq 3$ [3].

To clarify the statistics on this item, let N^+indicates to number of the positive values of the ranks determinants for size r, in addition to the zero value, then:

$$N^+ = r^{r-2} + 1 \dots \dots (2b)$$

That is clear, the number N^+ is equal to number of the absolute values of the Ranks Determinants.

And letN^*indicates to number of the positive values of the ranks determinants for size r, except zero, given by:

$$N^* = r^{r-2} \ \dots \dots (2c)$$

ISSN 1943-023X

Received: 16 July 2019/Accepted: 05 August 2019

The table (2.2) below shows (1a, 1b, 1c) numbers.

Table 2.2: The numbers of different

r	N	N^+	N^*
3	7	4	3
4	33	17	16
5	251	126	125

Ranks Determinant in terms r.

The n*th* term of an arithmetic sequence is given in form: $S_n = a + (n-1)d$ [4], and regarding the different absolute values of the Ranks Determinants, which are consists arithmetic sequence:$a = 0$is the first term, and$d$is referred to common difference. We can reformulate the n*th* term of the arithmetic sequence$S_n$by using the numbers: $Max|Det(A)|$,N^*, and d in term the generator variable r as follows:

$$Max|Det(A)|_r = N^* d \dots\dots \ (3)$$

The common difference$d$,in term the generator variable r, is given as the following:

$$\left.\begin{aligned} i)\ If\ r = 3\ , d = 6 \Rightarrow d = 2r \\ ii)\ If\ r = 4, d = 10 \Rightarrow d = \frac{5}{2} r \\ iii)\ If\ r = 5\ , d = 15 \Rightarrow d = 3r \end{aligned}\right\} \dots\dots (4)$$

Then, we'll got the arithmetic sequence of the common difference d:

$$2r, 2.5r, 3r, \dots . (\ finite)\ \dots\ \dots\ (5),$$

and its common difference is equal to$0.5r$.

In addition, by substitution formula.(2c), and substitution the terms of formula.(5) in formula.(3) is evaluated the formulas the maximum value of the $R_S D$in term the variable generator r as the following:

$$\left.\begin{aligned} i)\ Max|Det(A)|_{r=3} = 2r^{r-1} \\ ii)\ Max|Det(A)|_{r=4} = \frac{5}{2}\ r^{r-1} \\ iii)\ Max|Det(A)|_{r=5} = 3r^{r-1} \end{aligned}\right\} \dots\dots (6)$$

Then, the arithmetic sequence of the maximum values of the Ranks Determinant is:

$$Max|Det(A)|_r : 2r^{r-1}, 2.5\ r^{r-1}\ , 3r^{r-1} \dots (finite) \dots\dots (7)$$

and its common difference is equal to $0.5\ r^{r-1}$.

From formula.(5) the n*th* term is $\frac{1}{2}\ (r+1)$, and from .(7) the n*th* term is $\frac{1}{2}(r+1)r^{r-1}$, or :

$$Max|Det(A)|_r = \frac{1}{2}(r^r + r^{r-1}) \dots\dots (8)$$

The formula.(8) is general form of the maximum value of the Ranks Determinant in term the variable generator r.

III. The Relationship between the Value of the Ranks Determinant and the Correlation Amount between the Ranks

The distribution way of the ranks in columns of $R_S D$, plays a major role in producing the value of the $R_S D$.

On the other hand, the summation of the differences squares between the ranks for each two columns (denoted by $\sum d^2$), which are calculated by taken the combination of determinant columns, give an indication of the strength of the correlation of ranks and the value of the $R_S D$. The summation of the differences squares $\sum d^2$were found to be directly proportional to the absolute value of the $R_S D$. For example the table (3.1) shown these values when the size of the $R_S D$is $r = 3$, then one can show that $\sum d^2 \geq\ |Det(A)|$.

Table 3.1: The values of Ranks Determinants

$Det\ (A)$	-6	+6	-12	+12	-18	+18
$\sum d^2$	10	10	16	16	18	18

Received: 16 July 2019/Accepted: 05 August 2019

$r = 3$, and corresponding values of the summation Difference Squares between the ranks in column of Ranks Determinants.

In Table.(3.1), the value $Det(A) = 0$is excluded from this the observed conclusion, because some of them are containing at least two columns(rows) are equal, and the other is inexplicable in same cause, for example the $R_S DA = \begin{vmatrix} 2 & 4 & 1 & 4 \\ 4 & 2 & 4 & 1 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{vmatrix} = 0$.

That is clear:$|Det(A)| \propto \sum d^2$, and$M.C.C.R \propto \frac{1}{|Det(A)|}$, so the proportion$\frac{|Det(A)|}{Max\,|Det(A)|_r}$is an indicator of the relative strength of multiple correlation of ranks. And may be $|Det(A)| = Max|Det(A)|_r$.

IV. Conclusion Formula of Multiple Correlation Coefficient of Ranks

(4.1) Theorem: Let A is Ranks Determinant with size$r \geq 3$, r is finite positive number, and $|Det(A)|$ is the absolute value of the ranks determinant A, then the Multiple Coefficient Correlation of Ranks is given by the following formula:

$$M.C.C.R = 1 - \frac{|Det(A)|}{\frac{1}{2}(r^r + r^{r-1})} \quad \dots\dots (9)$$

Proof: for any size $r \geq 3$, r is finite positive number, the following inequality is true:

$- Max.Det(A) \leq Det(A) \leq Max.Det(A)$ Or; $-1 \leq \frac{Det(A)}{Max.Det(A)_r} \leq 1$. And ;$0 \leq \frac{|Det(A)|}{Max\,|Det(A)|_r} \leq 1$

$$\therefore pr\left(0 \leq \frac{|Det(A)|}{Max|Det(A)|_r} \leq 1\right)$$

It is clear that:$f(r) = \frac{|Det(A)|}{Max\,|Det(A)|_r}$, or ;$f(r) = \frac{|Det(A)|}{\frac{1}{2}(r^r+r^{r-1})}$ (form. (8))

The sigma field of probability F, is$0 \leq p(E) \leq 1$, $E \in F$, $p(s) = 1$[5][6] .i.e. $p + q = 1$,then$P = 1 - q$, $0 \leq q \leq 1$. So one can write: $q = \frac{|Det(A)|}{\frac{1}{2}(r^r+r^{r-1})}$, there for $P = M.C.C.R$, that is mean: $M.C.C.R = 1 - \frac{|Det(A)|}{\frac{1}{2}(r^r+r^{r-1})}$, proved.

V. Function of the Generator Variable

The function $f(r) = \frac{|Det(A)|}{Max\,|Det(A)|_r}$ is defined over the interval$[0, 1]$, so; $1 - f(r) = 1 - \frac{|Det(A)|}{Max\,|Det(A)|_r}$ is defined over same interval. Accordingly, the$M.C.C.R \in [0, 1]$. One canbe divided this interval into the following categories: *very weak*, *week*, *moderate*, *strong*, *very strong*. Alternatively, can divided in another categories.

In theorem. (4.1) is shown that $\frac{|Det(A)|}{Max\,|Det(A)|_r}$or $\frac{|Det(A)|}{\frac{1}{2}(r^r+r^{r-1})}$is function of *r*; this fact paves the following definitions. It is necessary to mention that the values of the $R_S D$ represented the differences between the ranks. If these differences increased, then the value of $R_S D$ will be increase too, and vice versa.

(5.1) *Definition*: The values of the Multiple Coefficient Correlation of Ranks given by the following formula:

$$M.C.CR = \begin{cases} +1, & Det(A) = 0 \\ 0, & |Det(A)| = Max|Det(A)|_r \\ (0, +1), & |Det(A)| < Max|Det(A)|_r \end{cases} \quad \dots\dots (9)$$

(5.2) *Definition*: by fact$f(r) = \frac{|Det(A)|}{Max\,|Det(A)|_r}$, one can represent of Multiple Coefficient Correlation of Ranks as the following function in r :

$$M.C.C.R = \begin{cases} +1, & f(r) = 0 \\ 0, & f(r) = 1 \\ (0, +1), & 0 < f(r) < 1 \end{cases} \quad \dots\dots (10)$$

Received: 16 July 2019/Accepted: 05 August 2019

VI. Examples

To illustrate how to apply the general formula (8) of the Multiple Correlation Coefficient of Ranks we put the following few examples.

Example (6.1)

Table 3.2: Ranks determinant 3×3

	Ranks of Option.1	Ranks of Option.2	Ranks of Option.3
Question.1	3	2	2
Question.2	2	3	1
Question.3	1	1	3

$Det(A) = 12$, $|Det(A)| = 12,$ $r = 3 \Rightarrow Max|Det(A)|_{r=3} = \frac{1}{2}(r^r + r^{r-1}) = 18$.. By (Form. (8)) we get: $M.C.C.R = 1 - \frac{12}{18} \Rightarrow M.C.C.R = 0.333$.

***Example* (6.2)**

Table 3.3: Ranks determinant 4×4

	Ranks of Option.1	Ranks of Option.2	Ranks of Option.3	Ranks of Option.4
Question.1	3	4	1	4
Question.2	2	3	2	1
Question.3	1	2	4	2
Question.4	4	1	3	3

$Det(A) = -100$, $|Det(A)| = 100,$ $r = 4 \Rightarrow Max|Det(A)|_{r=4} = \frac{1}{2}(r^r + r^{r-1}) = 160$. By (Form. (8)) we get:$M.C.C.R = 1 - \frac{100}{160} \Rightarrow M.C.C.R = 0.375$.

Example (6.3)

Table 3.4: Ranks determinant 5×5

	Ranks of Option.1	Ranks of Option.2	Ranks of Option.3	Ranks of Option.4	Ranks of Option.5
Question.1	4	2	3	3	1
Question.2	1	4	2	5	3
Question.3	3	1	4	2	5
Question.4	5	3	1	4	2
Question.5	2	5	5	1	4

$Det(A) = 1125$, $|Det(A)| = 1125,$ $r = 4 \Rightarrow Max|Det(A)|_{r=5} = \frac{1}{2}(r^r + r^{r-1}) = 1875$. By (Form. (7c)) we get:$M.C.C.R = 1 - \frac{1125}{1875} \Rightarrow M.C.C.R = 0.40$.

By same way we can calculate the M.C.C.R in $r = 6, 7, 8, 9, 10, \dots$. $r$is finite positive number.

References

[1] Yadolah Dodge (2008)" The Concise Encyclopedia of Statistics. ISBN: 978-0-387-32833, P. 500.

[2] Prasanna Sahoo(2013)" PROBANILITY AND MATHEMATICAL STATISCS", *USA*. P4.

[3] Abdulrahman. K. A. Zeyarah" Progressive Elements Determinant", *International Journal of Advance Research,* 6(12), 688-672, Vol 6, Issue 12, December 2018. ISSN: 2320-5407.

[4] CHARLES C. CARICO (1984) "COLLEGE ALGEBRA WITH ANALYTICAL GEOMETRY", *Printed in the United states of America.* ISBN 0-471- 88748-X, PP 196-198.

[5] B. C. Erricker(1981), "Advanced General Statistics" *HODDER ANDSTOGHT-ON* .

[6] Robert V. Hogg And Allen T .Craig(1970)" Introduction To Mathematical Statistics" – *THE UNIVERSTY of IOWA – Macmillan publishers LONDON - 3rd Edition.*

Received: 16 July 2019/Accepted: 05 August 2019