
The G-method for rank determination in rank order statistics for 'c' samples

Oyeka Ikewelugo Cyprian Anaene., Nwankwo Chike H., Awopeju K. Abidemi

Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria P.M.B 5025

Email address:

abidemiawopeju@gmail.com (A. K. Abidemi)

To cite this article:

Oyeka Ikewelugo Cyprian Anaene., Nwankwo Chike H., Awopeju K. Abidemi. The G-Method for Rank Determination in Rank Order Statistics for 'C' Samples. *American Journal of Theoretical and Applied Statistics*. Vol. 3, No. 1, 2014, pp. 18-24.

doi: 10.11648/j.ajtas.20140301.13

Abstract: This paper proposes a statistical method called 'the G method' to highlight its generalized nature for the determination and assignment of ranks to sample observations drawn from several populations for possible use in further analyses. The sampled populations may be measurements on as low as the ordinal scale and need not be continuous or even numeric. The proposed rank determination statistical model intrinsically and structurally provides for the breaking of possible ties between sample observations and automatically assigning such observations their mean ranks. This approach and hence the proposed model therefore obviate the need for the sampled populations to be continuous. They may be discrete or even non-numeric measurements on as low as the ordinal scale. The proposed method is of more generalized and wider applicability than an existing formulation which can be used with only continuous populations and is easier to use in practice than the usual traditional method which is often tedious and cumbersome, especially with large samples. The proposed method is illustrated with some data and shown to yield the same results as other existing methods where these methods are equally applicable.

Keywords: Tied Observation, Rank, Robust, Nonparametric, Rank-Order, G-Method

1. Introduction

A researcher may sometimes collect random samples drawn from some populations that are possibly measurements on as low as the ordinal scale and may prefer to use some more robust (nonparametric) rather than parametric methods for the data analysis. The researcher would in this situation need to first change the available raw scores into ranks for use for subsequent analyses. This is necessary because ranks and rank-order statistics are the basis for most nonparametric procedures providing the required data as functions of the original observations for use in statistical analyses. Rank order statistics for a random sample are any set of constants which indicate the order of the observations in terms of their magnitudes or relative relationships. The actual magnitude or size of any observation is used only in the determination of its relative position in comparison with other observations in the sample data set and thereafter ignored and not used in subsequent analyses based on rank-order statistics.

Rank-order statistics may alternatively be defined as the set of numbers which results when each original

observation [1] is replaced by the value of some order preserving function. Although theoretically any set of order preserving function may be used in the assignment of ranks to sample observations, in practice, for simplicity, only the set of the first positive integers, that is, a permutation of the first set of integers is preferably used for this purpose.

In assigning this set of ordered numbers or positive integers as ranks to sample observations the traditional and usual approach has often been to first arrange the sample observations either from the smallest to the largest or largest to the smallest and then assign them the positive integers as ranks accordingly, that is, either in increasing or decreasing order. This approach is however rather ad-hoc, heuristic and not systematized and formatted. Nevertheless an expression exists for the determination of the ranks that may be assigned to a set of sample observations [2]

The formula unfortunately can only be used mostly with populations that are continuous and numeric measurements. In continuous populations, no two observations are exactly equal to each other and the probability that any two

observations from such populations are exactly equal is theoretically zero. In such a situation the set of ranks obtained and assigned to sample observations using the currently existing formulae are always distinct positive integers. However in reality a set of observations do not always have different values. Some of the observations may have equal magnitudes or values and hence treated as tied observations.

Strictly speaking, the existing expression under inference cannot be used in breaking ties and determining the ranks of tied observations. Several methods however exist for breaking ties between sample observations in their ranking. If ties are few, the problem of tied observations may be resolved by dropping these tied observations and reducing sample sizes appropriately in subsequent analysis. The problem of ties, if they are not too many may also be resolved by assigning tied observations their mean ranks [2, 3, 4].

Thus, if there are only few ties and they could be dropped and ignored in subsequent analyses, then the currently existing expression for the determination of ranks for sample observations if they are numeric is still applicable. If however the ties are not few and there is a need to break tied observations by assigning them their mean ranks, then the existing formulation may not be readily used for this purpose. Furthermore, if the sampled populations are non-numeric measurements on as low as the ordinal scale, such as letter grades or scores, then the currently existing expression is of no use in assigning ranks to the observations. Also if several samples are to be

Let

$$\mu_{ij} = \begin{cases} 1, & \text{if the } j^{\text{th}} \text{ observation or score } x_j \text{ is at least as high (good, large, great, serious) as} \\ & \text{the } i^{\text{th}} \text{ observation or score } x_i \text{ or } x_j \geq x_i \\ 0, & \text{if the } j^{\text{th}} \text{ observation or score } x_j \text{ is lower (worse, smaller, less, less serious)} \\ & \text{than the } i^{\text{th}} \text{ observation or score } x_i, \text{ or } x_j < x_i. \end{cases} \quad (1)$$

for $i, j = 1, 2 \dots n$.

Note that μ_{ij} of eqn. 1 is defined and applicable to all data sets whether or not continuous and whether or not numeric provided they are measurements on at least the ordinal scale.

Note also that eqn. 1 does not require that the sample observations or scores being analyzed be presented or arranged in any predetermined or pre-specified form or order.

Now using Eqn. 1, we have that the rank order statistic $r(x_j) = r_j$, that is the rank assigned to the j^{th} observation or score in the ranking of the 'n' sample observations from the largest or highest to the smallest or least may be determined from the expression

$$r(x_j) = r_j = \sum_{i=1}^n \mu_{ij} \quad (2)$$

for some $j = 1, 2 \dots n$

combined and ranked together then no matter the nature of the sampled populations, use of either the traditional method or the currently existing formulation to rank the observations would be tedious, laborious or impossible.

In this paper, we propose to develop a more generalized formulation or method that would enable a researcher to more systematically assign ranks to random samples drawn from several populations measured on at least the ordinal scale whether or not the populations are continuous or discrete and whether or not they are numeric or non-numeric.

2. The Proposed Generalized Method (G-Method)

Suppose x_{il} is the i^{th} observation or score in a random sample of size n_l drawn from population x_l for $i = 1, 2, \dots, n_l$; $l = 1, 2, \dots, c$. Population x_l may be measurements on as low as the ordinal scale and need not be continuous or numeric. Research interest is to develop a statistical method or procedure to systematically assign ranks to observations in each of the 'c' samples when all the $n = \sum n_l$ sample observations are pooled and assigned ranks for use in further statistical analysis.

To do this, we may pool the 'c' samples together in any desired way or as presented. Let x_i be the i^{th} observation in the pooled sample, that is, the 'c' samples combined for $i = 1, 2 \dots n$.

Equation 2 enables the researcher assign unique ranks to the 'n' sample observations or scores, if these observations are all distinct. However the n sample observations may not always all be distinct. They may consist of some K groups of different observations with each group containing observations that have equal values and hence treated as tied observations. Thus in the situation the rank r_j of the j^{th} observation would then in effect be the rank r_h common to some t_h observations tied in value in the h^{th} group of tied observation for $h = 1, 2 \dots k$.

This means that the j^{th} observation x_j which now also may be taken as the h^{th} observation in the i^{th} group of tied observations is tied in score with t_{h-1} other observations in that group. Research interest may therefore be in resolving the problem of ties between sample observations possibly by assigning mean ranks to observations tied in value.

Now, suppose assigning mean ranks is the preferred approach to resolving the problem of ties between tied observations, suppose also that x_j assigned the rank $r(x_j) =$

r_j using equation 2 is tied in value with some t_{h-1} other observations or scores in the h^{th} group of tied observations in the sample and hence is now also treated as, that is regarded as r_h the rank common to the h^{th} group of tied observation, for $h = 1, 2, \dots, K$, then, the mean or average rank $\text{avr}(r_h; t_h)$ of t_h tied observations in the h^{th} group of tied observations that is for the h^{th} observation, x_h tied in value with some t_{h-1} other observations in the h^{th} group of tied observations is determined from the expression.

$$\text{avr}(r_h, t_h) = r_h - \frac{(t_h - 1)}{2} \quad (3)$$

Note that if $t_h = 1$, that is, if the h^{th} observation with rank r_h is not tied in value with any other observation in the h^{th} group of tied observations in the sample, then, $r_j = r_h$ for $j = 1, 2, \dots, n$ and some $h = 1, 2, \dots, k$.

The sum R_s of the assigned ranks of a given sample s is easily calculated from the results of eqn. 3 for samples as the weighted sum of the average ranks $\text{avr}(r_{hs}; t_{hs})$ of the ranks \bar{r}_{hs} assigned to each of the t_{hs} observations tied in value in the h^{th} group of tied observations in sample s for $h = 1, 2, \dots, k; s = 1, 2, \dots, c$, that is

$$\begin{aligned} R_s &= \sum_{h=1}^k t_{hs} \cdot \text{avr}(r_{hs}, t_{hs}) \\ &= \sum_{h=1}^k t_{hs} \cdot r_{hs} - \frac{1}{2} \sum_{h=1}^k (t_{hs}^2 - t_{hs}) \end{aligned} \quad (4)$$

for $s = 1, 2, \dots, c$

The mean of the ranks \bar{r}_s assigned to sample 's' is then calculated as R_s divided by $\sum_{h=1}^k t_{hs} = n_s$ the size of sample 's'.

That is

$$\bar{r}_s = \frac{R_s}{\sum_{h=1}^k t_{hs}} = \frac{\sum_{h=1}^k t_{hs} \cdot r_{hs} - \frac{1}{2} \sum_{h=1}^k (t_{hs}^2 - t_{hs})}{\sum_{h=1}^k t_{hs}} \quad (5)$$

Where, $\sum_{h=1}^k t_{hs} = n_s$, for $s = 1, 2, \dots, c$

Similarly the total sum of the ranks assigned to the 'c' samples together is from eqns. 3 and 4

$$\begin{aligned} R &= \sum_{h=1}^k t_{hs} \cdot \text{avr}(r_h, t_h) \\ &= \sum_{h=1}^k t_h \cdot r_h - \frac{1}{2} \sum_{h=1}^k (t_h^2 - t_h) \\ &= \sum_{s=1}^c \sum_{h=1}^k t_{hs} \cdot \text{avr}(r_{hs}, t_{hs}) \\ &= \sum_{s=1}^c \sum_{h=1}^k t_{hs} \cdot r_{hs} - \frac{1}{2} \sum_{s=1}^c \sum_{h=1}^k (t_{hs}^2 - t_{hs}) \end{aligned} \quad (6)$$

Note that $R = \frac{n(n+1)}{2}$ the total sum of ranks assigned to

the total of 'n' observation in the combined 'c' samples.

The corresponding overall mean-rank 'r' for 'n' sample observations is

$$\bar{r} = \frac{R}{\sum_{s=1}^c \sum_{h=1}^k n_{hs}} = \frac{\sum_{h=1}^k t_h \cdot r_h - \frac{1}{2} \sum_{h=1}^k (t_h^2 - t_h)}{\sum_{s=1}^c \sum_{h=1}^k n_{hs}} \quad (7)$$

where $\sum_{s=1}^c \sum_{h=1}^k n_{hs} = \sum_{s=1}^c n_s = n$ the overall sample size.

Equations 3-7 would readily enable the researcher determine the ranks and mean or average ranks to be assigned to each observation in each of the 'c' samples separately when these samples are combined and ranked together as one common sample as well as the total and mean ranks both overall and for each sample. Thus, the method formulations are directly used in calculations to provide necessary inputs used in hypothesis testing, a procedure that is likely to be quicker, more helpful and cost effective than the usual traditional approach.

3. Illustrative Example

A certain university has three possible modes of study for her students namely full time with class contact, part time with class contact or electronic (e) learning without class-contact. A course instructor returned the following scores in letter grades for students who took his course offered in any of the three study modes during the semester.

Full Time: A⁻, E, C, B⁺, A, D, F, B⁻, A⁺, B⁻, B, C⁺, A⁺,

Part Time F, F, B⁺, D, A⁻, A, D, C⁺ A⁺

e-learning: F, A⁻, A, A, E, A⁺, B⁺, A, A⁻, A, A⁺

Research interest is to determine whether students perform equally well in course under the three modes of study in the university. If the researcher wishes to test the null hypothesis of equal performance by students under the three modes of study using an appropriate non parametric test based on ranks, then the researcher would first need to convert the above letter grades into ranks for use in data analysis.

To do this using the proposed method we may pool the three samples of letter grades in any form into one combined sample of size $n = 33$ and then apply eqn. 1. However, for simplicity, the samples would here be pooled and listed according to mode of study from full time to e-learning and then apply eqn. 1 to obtain the values of 1s and 0s of μ_{ij} , for $i, j = 1, 2, \dots, 33$. The values of μ_{ij} may, for greater clarity, be presented in tabular form as in Table 1.

The column totals of Table 1 namely; $r(x_j) = r_j$ are the ranks of the letter grades ranked from the poorest F assigned the lowest rank to the best grade A⁺ assigned the

highest rank. Similarly, the row totals namely $r(x_i) = r_i$ are the ranks of the letter grades ranked from the best grade A⁺ assigned the smallest rank to the worst grade F assigned the highest rank in the combined ranking of the 33 letter grades.

Using these total column ranks $r(x_j) = r_j$ in equation 3 we obtain the ranks and average ranks, $avr (r_h; t_h)$ for the various letter grades. The results are presented in Table 2.

Table 1. Values of μ_{ij} (Eqn 1) in the combined Ranking of grades under three study mode

S/N	Letter (xi)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		A ⁻	E	C	B ⁺	A	D	F	B ⁻	A ⁺	B ⁻	B	C ⁺	A ⁺	F	F
1	A ⁻	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0
2	E	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0
3	C	1	0	1	1	1	0	0	1	1	1	1	1	1	0	0
4	B ⁺	1	0	0	1	1	0	0	0	1	0	0	0	1	0	0
5	A	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0
6	D	1	0	1	1	1	1	0	1	1	1	1	1	1	0	0
7	F	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	B ⁻	1	0	0	1	1	0	0	1	1	1	1	0	1	0	0
9	A ⁺	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
10	B ⁻	1	0	0	1	1	0	0	1	1	1	1	0	1	0	0
11	B	1	0	0	1	1	0	0	0	1	0	1	0	1	0	0
12	C ⁺	1	0	0	1	1	0	0	1	1	1	1	1	1	0	0
13	A ⁺	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
14	F	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	F	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	B ⁺	1	0	0	1	1	0	0	0	1	0	0	0	1	0	0
17	D	1	0	1	1	1	1	0	1	1	1	1	1	1	0	0
18	A ⁻	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0
19	A	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0
20	D	1	0	1	1	1	1	0	1	1	1	1	1	1	0	0
21	C ⁺	1	0	0	1	1	0	0	1	1	1	1	1	1	0	0
22	A ⁺	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
23	F	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
24	A ⁻	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0
25	A	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0
26	A	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0
27	E	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0
28	A ⁺	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
29	B ⁺	1	0	0	1	1	0	0	0	1	0	0	0	1	0	0
30	A	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0
31	A ⁻	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0
32	A	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0
33	A ⁺	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
Total		22	6	10	18	28	9	4	14	33	14	15	12	33	4	4

Table 1. continue

16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
B⁺	D	A⁻	A	D	C⁺	A⁺	F	A⁻	A	A	E	A⁺	B⁺	A	A⁻	A	A⁺
0	0	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1	1
1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
1	0	1	1	0	1	1	0	1	1	1	0	1	1	1	1	1	1
1	0	1	1	0	0	1	0	1	1	1	0	1	1	1	1	1	1
0	0	0	1	0	0	1	0	0	1	1	0	1	0	1	0	1	1
1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	0	1	1	0	0	1	0	1	1	1	0	1	1	1	1	1	1
0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1
1	0	1	1	0	0	1	0	1	1	1	0	1	1	1	1	1	1
1	0	1	1	0	0	1	0	1	1	1	0	1	1	1	1	1	1
1	0	1	1	0	1	1	0	1	1	1	0	1	1	1	1	1	1
0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	0	1	1	0	0	1	0	1	1	1	0	1	1	1	1	1	1
1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1
0	0	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1	1
0	0	0	1	0	0	1	0	0	1	1	0	1	0	1	0	1	1
1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1
1	0	1	1	0	1	1	0	1	1	1	0	1	1	1	1	1	1
0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1	1
0	0	0	1	0	0	1	0	0	1	1	0	1	0	1	0	1	1
0	0	0	1	0	0	1	0	0	1	1	0	1	0	1	0	1	1
1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1
1	0	1	1	0	0	1	0	1	1	1	0	1	1	1	1	1	1
0	0	0	1	0	0	1	0	0	1	1	0	1	0	1	0	1	1
0	0	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1	1
0	0	0	1	0	0	1	0	0	1	1	0	1	0	1	0	1	1
0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1
18	9	22	28	9	12	33	4	22	28	28	6	33	18	28	22	28	33

Table 2. Ranks and Average Ranks $avr(r_h; t_h)$ (Eqn 3) from Table 1 for the combined sample of letter grades

Group (h)	Scores letter Grades (x_i)	Group Rank (r_h)	No of grades in group (t_h)	Rank of Average Rank $(avr(r_h; t_h)) = r_h - \left(\frac{t_h - 1}{2}\right)$	Weighted Average Rank $(t_h \cdot avr(r_h; t_h))$
1	A ⁻	22	4	20.5	82
2	E	6	2	5.5	11
3	C	10	1	10	10
4	B ⁺	18	3	17	51
5	A	28	6	25.5	153
6	D	9	3	8	24
7	F	4	4	2.5	10
8	B ⁻	14	2	13.5	27
9	A ⁺	33	5	31	155
10	B	15	1	15	15
11	C ⁺	12	2	11.5	23
Overall Total Rank (R)					561
Overall Total Mean Rank = \bar{r}					17.00

Table 3. Ranks and Mean Ranks of letter grades in each of the three samples in their combined ranking (Eqn 3)

Grade	Full time students		Grade	Part time student		Grade	e-learning students	
	Rank (Eqn 3 Table 2)	Traditional ranking method		Rank (Eqn 3 Table 2)	Traditional Ranking Method		Rank (Eqn 3 Table 2)	Traditional Ranking Method
A ⁻	20.5	20.5	F	2.5	2.5	F	2.5	2.5
E	5.5	5.5	F	2.5	2.5	A ⁻	20.5	20.5
C	10	10	B ⁺	17	17	A	25.5	25.5
B ⁺	17	17	D	8	8	A	25.5	25.5
A	25.5	25.5	A ⁻	20.5	20.5	E	5.5	5.5
D	8	8	A	25.5	25.5	A ⁺	31	31
F	2.5	2.5	D	8	8	B ⁺	17	17
B ⁻	13.5	13.5	C ⁺	11.5	11.5	A	25.5	25.5
A ⁺	31	31	A ⁺	31	31	A ⁻	20.5	20.5
B ⁻	13.5	13.5	Sample Total R_s	126.5		A	25.5	25.5
B	15	15	Sample Mean = \bar{r}	14.06		A ⁺	31	31
C ⁺	11.5	11.5			Sample Total R_s	230		
A ⁺	31	31			Sample Mean = \bar{r}	20.91		
Total Rank (R)	204.5							
Mean Rank = \bar{r}	15.73							

Now, from table 2 and consistent with eqn. 3, we obtain the ranks or mean ranks of each letter grade for each of the three samples of letter grades or scores in their combined ranking from the poorest grade F to the best grade A⁺. The results are shown in Table 3 together with the corresponding ranks that would have been obtained if the scores had been ranked using the traditional or usual method.

Notice from Table 3 that the proposed method yield

exactly the same rank for each observation (grade) in the combined ranking of the same observation as would be obtained using the traditional method of ranking of observations.

Use of eqn. 6 with the results of Table 2 gives the overall total sum of the ranks as shown in the last column of Table 2 as:

$$R = (4) (20.5) + (2) (5.5) + \dots + (2)(11.5) = 561.$$

Similarly the overall mean rank is from eqn. 6 and table 2

$$\bar{r} = \frac{561}{33} = 17.00$$

The corresponding sample total ranks and mean ranks are similarly calculated using eqn. 4 and 5 respectively. The results are shown in Table 3

Notice that as expected

$$R_1 + R_2 + R_3 = 204.5 + 126.5 + 230 = 561 = R.$$

The overall total mean rank

$$\frac{n_1 \bar{r}_1 + n_2 \bar{r}_2 + n_3 \bar{r}_3}{n}$$

$$= \frac{(13)(15.73) + (9)(14.06) + (11)(20.9)}{33} = \frac{561.04}{33} = 17.00$$

These are the same results that would have been obtained using the traditional method or other applicable methods for determining ranks for sample observations. However the method proposed and presented here is more systematic and statistically formulated and formatted and easily enables a fuller and clearer understanding of data ranking process than would the traditional method.

4. Summary and Conclusion

We have, in this paper proposed and presented a statistical method for determining and assigning ranks to sample observations in random samples drawn from

several populations that may be measurements on as low as the ordinal scale. The sampled populations need not be continuous or even numeric.

The results obtained using the proposed method are the same as would have been obtained using the traditional approach and other existing methods if equally applicable. However, the present method is more generalized, systematic and formatted. It easily enables a fuller and clearer understanding of data ranking process than would the traditional method. Also, the proposed method, unlike is often the case, with the traditional approach does not require the original sample data to be first arranged in some form such as from the smallest to the largest or largest to the smallest before their ranking which is an additional advantage of the present method. The proposed method is easier, quicker and more cost effective to use especially when large samples with many tied observations are to be combined and ranked together for use in further analyses.

References

- [1] Wikipedia (November, 2013): Ranking the tied encyclopedia.
- [2] Gibbons, J.D. (1973): Non- Parametric Statistics. An Introduction; Newbury Park: Sage Publication
- [3] Hollander, M. and Wolfe, D.A. (1999): Non-Parametric Statistical Methods, (2nd Edition). Wiley Inter-science, New York
- [4] Siegel S. (1956): Nonparametric Statistics for the Behavioural Sciences. Int. Student Edition. McGraw-Hill Kogakusha Ltd, Tokyo.