SOME COMBINATORIAL PROBLEMS RELATED TO RANK SUMS IN A TWO - WAY CLASSIFICATION

FUAT ÖZKAN

Department of Statistics, Middle East Technical University Ankara, Turkey

ABSTRACT

Statistical problems may sometimes require the use of ranks perhaps due to the reason that the response can only be ranked but not mesasured. Yet in some other problems, the utilization of ranked data may be a desirable device to avoid an unjustified assumption of normality, implicit in the Analysis of Variance.

However, the analysis of ranked data may in turn lead to some complicated combinatioral problems. The author discusses some of these problems and presents theorems and formulations.

INTRODUCTION

The usual approach to statistical problems is to use regression analysis or analysis of variance which in turn involve the application of normal theory methods such as t-tests and F-tests. Frequently however, the assumption of normality for the random variables are not justified. Many types of industrial, economic or social data are nonnormal.

When the data are sufficiently extensive to indicate nonnormality, it might be desirable to use non–parametric methods where, instead of using the original quantitative values, ranks might be assigned to measurements. In this way no assumption whatever is required for the distribution of the random variables. The application of ranks to measurements may thus be a desirable device to avoid normality assumptions.

Furthermore, the use of ranked data may sometimes be inescapable because the random variable is a qualitative characteristic which can be ranked but not measured.

This situation has consequently forced researchers to develop nonparametric inferential methods based upon ranks. Friedman (1937) outlined a procedure whereby the analysis of ranked data could be employed in place of the ordinary analysis of variance when there were two or more criteria of classification. Friedman's test used the ranks within blocks to test for the main effects of a single factor. As Mehra and Sen (1969) showed this procedure causes information about interblock differences to be lost.

Hora and Conover (1984) demonstrated that the limiting null distribution of the usual F-statistics for the main efects in the two–way layout had the same limiting distribution when applied to ranks. Their procedure was to rank all treatments simultaneously without regard to block membership or level of treatment. The usual para-metric analysis of variance was then applied to the ranks.

Recently many other publications appeared in this area. Campbell and Skillings (1985) discussed a nonparametric multiple comparison with particular emphasis given to stepwise procedures.

Cuzick (1985) studied asymptotic properties of censored linear rank tests. Max (1985) gave an interpretation to the expected ranks of k objects ordered by randomly generated observers.

Acar and Pettitt (1985) introduced a technique based on ranks of observations which used an approximation to a marginal likelihood of ranks to find predictive probabilities for a future response.

Salter and Fawcett (1985) studied small sample robustness and power of an aligned rank transformation statistics, employing Monte Carlo methods. Schluchter (1985) extended the aligned rank test of Hodges and Lehmann (1962) to censored survival data collected in matched pairs or randomized blocks.

Henery (1986) started with the average ranks allocated by m judges to k objects and used order statistics models to find a value for the average of Kendall's (1970) τ between the judges' rankings and the true ranking.

Almost all of the papers mentioned above discuss procedures that require a good background in statistics and a considerable amount of computations. To overcome this obstacle, the author has conducted studies in the development of some quick and easy to use significance tests based upon ranks sums.

This paper aims at discussing only some of the combinatorial problems involved in a rank sum test in a two way classification.

However, to lay the necessary ground work for the analysis, first a description of the rank sum test will be given and then the combinatorial problems analyzed. For consideration's leading to the developments in this study and complementary details, the reader is referred to Özkan's (1975) previous work on this topic.

COMBINATORIAL PROBLEMS RELATED TO RANK SUMS IN A TWO–WAY LAYOUT

As stated in the introduction, the data for the problem under consideration may originate from an experiment where the response can only be ranked but not measured. In another instance the original data is in the form of measurements but the normal theory methods can not be applied because, assumptions of independence of observations, constancy of variance and normality can not be justified.

To show the basic structure of the problem let us consider the following two–way classification, where all observations are ranked simultaneously, using positive integers from 1 to pb, without regards to the membership to the first or the second factors.

Table 1. The Structure of a Two-way Layout

	Levels of second factor				Totals
Tevels of first factor	x_1	x_2	x_j	x_b	
z_1	x_{11}	x_{12}	x_{1j}	x_{1b}	U_1
z_2	x_{21}	x_{22}	x_{2j}	x_{2b}	U_2
.	.	.	.	.	.
.	.	.	.	.	.
z_i	x_{i1}	x_{i2}	x_{ij}	x_{ib}	U_i
.	.	.	.	.	.
.	.	.	.	.	.
z_p	x_{p1}	x_{p2}	x_{pj}	x_{pb}	U_p
Totals	T_1	T_1	Tj	T_b	

The first and the second factors have levels $Z_1, Z_2, \ldots, Z_p$ and $X_1, X_2 \ldots, X_b$ respectively. The x_{ij} in the table shows the rank chosen from positive integers from 1 to pb and assigned to the response for the Z_ith level of the first factor and X_jth level of the second factor. U_i is the sum of ranks for the Z_ith level of the first factor. T_j is the sum

of ranks assigned to the X_j th level of the second factor. As an example Z_i might designate the treatments and X_j might show the blocks in a randomized block design.

It is possible to develop a slippage test based upon the rank sums U_i and T_j and test the hypothesis H_0, which states that the rankings are chosen at random from the collection of all permutation of numbers, $1, \ldots, pb$ and they are independent. The test presented will be powerful especially against the alternatives that one level of a factor has larger probability than the other ones of being ranked high (or low), whilst the other $(b-1)$ [or $(p-1)$] levels are ranked in a random order. For the first factor $H_0: U_1 = U_2 = \ldots = Up$ will be tested against the p alternatives that Ha: U_i is larger than the others, $(i = 1, 2, \ldots, p)$ or it will be tested against the alternatives that Ha: U_i are smaller than the others $(i = 1, 2, \ldots, p)$.

For the second factor, $H_0: T_1 = T_2 = \ldots = T_b$ will be tested against the alternative that $H_a: T_j$ $(j = 1, 2, \ldots, b)$ is larger that the others or H_0 will be tested against the alternative that Ha: T_j $(j = 1, 2, \ldots, b)$ is smaller than the others.

In this paper as stated before, we will only study the combinatorial problems related to U_i and T_j.

For U_i we have,

$$\frac{b(b+1)}{2} \leq U_i \leq \frac{(2bp - b + 1)b}{2} \tag{1}$$

Since for the smallest value of U_i, the first b ranks; from the set of ranks ordered in the increasing order of magnitiude, must be summed. And for the largest U_i, the last b ranks, from this set would have to be added up. Thus if $U_{max.}$ designates the largest possible value of U.

$$U_{max.} = (pb - b + 1) + (pb - b + 2) + \ldots + [pb - b + (b-1)]$$
$$+ pb = \frac{b}{2}(2pb - b + 1) \tag{2}$$

And,

$$\sum_{i=1}^{p} U_i = \frac{bp(pb+1)}{2} \tag{3}$$

which is the sum of numbers from 1 to pb. The average of U_i

$$\bar{U} = \frac{pb(bp+1)}{2p} = \frac{b(pb+1)}{2} \tag{4}$$

We also have bounds for the totals of T_j, T_1, T_2, ..., T_b,

$$\frac{p(p+1)}{2} \leq T_j \leq \frac{(2pb - p + 1)p}{2} \tag{5}$$

The smallest T_j is formed by the first p ranks and the largest T_j is the total of the last p ranks, in the ordered set.

$$T_{max.} = (pb - p + 1) + (pb - p + 2) + ... + [pb - p + (p - 1)] + pb$$

$$= \frac{p}{2}(2pb - p + 1) \tag{6}$$

$$\sum_{j=1}^{p} T_j = \frac{pb(pb+1)}{2} \tag{7}$$

$$\overline{T} = \frac{pb(bp+1)}{2b} = \frac{p(bp+1)}{2} \tag{8}$$

The number of ways U_i can be formed is given by

$$C_b^{pb} = \frac{(pb)!}{b!(pb-b)!} \tag{9}$$

For example for a randomized block design consisting of two blocks and three treatments, the rank totals for blocks can be formed in $C_3^6 = 20$ ways and they are, as in Table 1.:

T_j	:	6	7	8	9	10	11	12	13	14	15
Frequency	:	1	1	2	3	3	3	3	2	1	1

We can express this in a combinatorial generating function form that is named f (6–3), which shows all possible rank sums of three ranks, taken from 1, 2, ..., 6. To express f (6–3) we first write the possible combinations of six ranks taken three at a time and then show the totals as powers of a symbol, say (a). Thus we have,

$$f(6-3) = a^6 + a^7 + a^8 + a^9 + a^9 + a^{10} + a^{11} + a^{12} + a^{13}$$

$$+ a^{15} + a^8 + a^9 + a^{10} + a^{10} + a^{11} + a^{12} + a^{11} + a^{12}$$

$$+ a^{13} + a^{14} \tag{10}$$

(10) can be put in a more compact form,

$$f(6-3) = a^6(1 + a^2 + a^3 + a^5)(1 + a + a^2 + a^3 + a^4) \tag{11}$$

As other examples we can write

$$f(6-2) = a^3 (1 + a + a^2 + a^3 + a^4)(1 + a^2 + a^4) \tag{12}$$

$$f(7-2) = a^3 (1 + a + a^2 + a^3 + a^4 + a^5 + a^6)(1 + a^2 + a^4) \tag{13}$$

For the general combinatorial problem of n ranks, 1, ..., n, taken two at a time we state a theorem

Table 1. C^6_3 Combinations, T_j sums and symbols.

Combinations			Sums (T_j)	Symbols		
1	2	3	6	$a\ a^2 a^3$	or	a^6
1	2	4	7	$a\ a^2 a^4$	"	a^7
1	2	5	8	$a\ a^2 a^5$	"	a^8
1	2	6	9	$a\ a^2 a^6$	"	a^9
2	3	4	9	$a^2 a^3 a^4$	"	a^9
2	3	5	10	$a^2 a^3 a^5$	"	a^{10}
2	3	6	11	$a^2 a^3 a^6$	"	a^{11}
3	4	5	12	$a^3 a^4 a^5$	"	a^{12}
3	4	6	13	$a^3 a^4 a^6$	"	a^{13}
4	5	6	15	$a^4 a^5 a^6$	"	a^{15}
1	3	4	8	$a^1 a^3 a^4$	"	a^8
1	3	5	9	$a^1 a^3 a^5$	"	a^9
1	3	6	10	$a^1 a^3 a^6$	"	a^{10}
1	4	5	10	$a^1 a^4 a^5$	"	a^{10}
1	4	6	11	$a^1 a^4 a^6$	"	a^{11}
1	5	6	12	$a^1 a^5 a^6$	"	a^{12}
2	4	5	11	$a^2 a^4 a^5$	"	a^{11}
2	4	6	12	$a^2 a^4 a^6$	"	a^{12}
2	5	6	13	$a^2 a^5 a^6$	"	a^{13}
3	5	6	14	$a^3 a^5 a^6$	"	a^{14}

Theorem 1.

The rank sums for all possible combinations of ranks 1, 2, 3, ..., n, taken two at a time are given by the combinatorial generating function shows below in (14).

$$f(n-2) = a^3 \left[1 + a + a^2 + a^3 + \ldots + a^{\frac{2n-3-(-1)^n}{2}} \right] \left[1 + a^2 \right.$$

$$\left. + a^4 + a^6 + \ldots + a^{\frac{2n-5-(-1)^{n-1}}{2}} \right] \tag{14}$$

Before proving theorem 1 let us examine the right side of (14). Note that the first bracket contains terms which are multiples of (a), as 1, a, a^2, a^3. ... To find the last term in the first bracket we must determine the power of the last term given by

$$[2n-3-(-1)^n] / 2 \tag{15}$$

If (15) is zero this means that there is only one term in the first bracket and it is equal to 1.

For $n = 2$, (15) becomes zero, but it can not take negative values, for we must have $n \geq 2$, which means that to form combinations of two ranks, we must have at least two ranks to begin with. Thus in general $2\,n \geq 4$ and

$$2n-3-(-1)_n \geq 0 \tag{16}$$

To show (16) let n be even, then $(-1)^n = 1$ and consequently.

$$2n-3-1 = 2n-4 \geq 0$$

If n is odd $(-1)^n = -1$ and we have $2n-3 + 1 = 2n - 2 \geq 0$.

Thus we have shown that for $n \geq 2$ there will be at least one term in the first bracket of (14) and it is 1.

If (15) takes a value other than zero, say k, the first bracket of (14) will take the form $[1 + a + a^2 + a^3 + \ldots + a^k]$ and it will contain $(k + 1)$ terms.

The second bracket on the right side of (14) contains a sequence of terms which are multiples of a^2, as, 1, a^2, a^4, a^6, ... To find the last term in this sequence, we examine the power of the last term

$$[2n-5-(-1)^{n-1}]\,/\,2 \tag{17}$$

If (17) is zero this means there is only one term in the second bracket of (14) and it is (1).

The value of (17) becomes zero for the smallest meaningful value of $n = 2$. Hence for $n = 2$ (14) takes the form: $f\,(2-2) = a^3\,(1)\,(1)$ $= a^3$, which shows that there is only one combination of ranks 1 and 2 and their sum is equal to 3. This proves (14) for $n = 2$. (17) can not become negative since, $n \geq 2$ and for n odd, $n = 3, 5, 7, \ldots$

$$2n-5-1 \geq 0$$

For $n \geq 2$ and even, (17) takes the form,

$$2n-5 + 1 = 2n-4 > 0$$

There are $[2n-1-(-1)^n]\,/\,2$ terms in the first bracket of (14).

Since, its terms start with 1 and continue as a, a^2, a^3, ..., thus the sequence contain one more term than the power of the last term. That is

$$[2n-3-(-1)^n]\,/\,2 + 1 = [2n-1-(-1)^n]\,/\,2 \tag{18}$$

The number of terms in the second bracket is $[2n-1-(-1)^{n-1}]\,/\,4$. This is because its terms are in the form $1 + a^2 + a^4 + a^6 + \ldots$ and their number is equeal to the half of the power of the last term plus 1.

$(1 / 2) [2n–5–(–1)^{n-1}] / 2 + 1 = [2n–1–(–1)^{n-1}] / 4$

We will now prove Theorem (1) by the method of induction.

1. We have already shown that (14) is true for the smallest meaningful value of n = 2. Next,

2. By assuming the expression shown in (14) to be true for n we have to show that it is also true for (n + 1).

Let's assume that (14) is true for n. Next we increase n by 1 and examine the additional combinations of two ranks caused by this change. The new enlarged set of ranks from which the combinations of two will be formed is $1, 2, \ldots, n, (n + 1)$. The combinations of this set taken two at a time, will include the combinations of the elements of $1, 2, \ldots, n$ taken two at a time plus the (n) combinations of two of the rank (n + 1) and all the other elements from 1 to n. Thus the additional combinations and their symbols are shown below:

$$[1, (n + 1)], [2, (n + 1)], [3, (n + 1)] \ldots [n, (n + 1)] \qquad (19)$$

$$a^{n+2}, a^{n+3}, a^{n+4}, \ldots, a^{2n+1} \qquad (20)$$

Hence,

$$f[(n + 1)-2] = f(n-2) + \text{the elements in (20)} \qquad (21)$$

(20) can be written as,

$$a^3 \cdot a^{n-1} (1 + a + a^2 + \ldots + a^{n-1}) \qquad (22)$$

Now (21) becomes),

$$f[(n + 1)-2] = a^3 \underbrace{[(1 + a + a^2 + \ldots + a^{\frac{2n-3-(-1)^n}{2}}}_{**}) (1 + a^2 + a^4$$

$$+ a^6 + \ldots + a^{\frac{2n-5-(-1)^{n-1}}{2}}) + a^{n-1} \underbrace{(1 + a + a^2 + \ldots + a^{n-1})}_{*}] \qquad (23)$$

For n, odd

$$a^{\frac{2n-3-(-1)^n}{2}} = a^{\frac{2n-3+1)}{2}} = a^{n-1} \qquad (24)$$

Thus in (23) the expressions shown by (**) and (*)

$$1 + a + a^2 + \ldots + a^{\frac{2n-3-(-1)^n}{2}} \qquad (25)$$

and

$$1 + a + a^2 + \ldots + a^{n-1} \tag{26}$$

become equal, that is (25) takes the form

$$1 + a + a + a^2 + \ldots + a^{n-1} \tag{27}$$

Furthermore for n odd

$$[2n-5- (-1)^{n-1}] / 2 = n-3$$

Thus (23) can be written as,

$$f[(n + 1)-2] = a^3 (1 + a + a^2 + \ldots + a^{n-1}).$$

$$(1 + a^2 + a^4 + \ldots + a^{n-3} + a^{n-1}) \tag{28}$$

We substitute (n + 1) for n in (14) to see its form for (n + 1) ranks. First we calculate the following powers, where (n + 1) is even, since n is odd.

$$[2 (n + 1) -3- (-1)^{n+1}] / 2 = n-1 \tag{29}$$

$$[2 (n + 1) -5- (-1)^{n}] / 2 = n-1 \tag{30}$$

Hence (14) becomes identically equal to (28) and this proves the truth of (14) for n odd. Now we have to prove it for n even. For n even (15) becomes (n–2) and (17) also becomes (n–2). Therefore (14). the truth of which we assume for (n) takes the form,

$$f(n-2)=a^3 (1+a+a^2 +\ldots+ a^{n-2}) (1+a^2+a^4 +\ldots+ a^{n-2}) \tag{31}$$

Now examining (22) which shows the additional tems caused by going from n to (n + 1), we note that the number of terms is equal to n, and by assumption n is even. We can therefore show the terms in (22) in two groups, each having n / 2 terms, as

$$a^3. a^{n-1} \underbrace{(1 + a^2 + a^4 + \ldots + a^{n-2}}_{n/2 \text{ terms}} \underbrace{+ a + a^3 + a^5 + \ldots + a^{n-1})}_{n/2 \text{ terms}} \tag{32}$$

We can also express (32) as

$$a^3 (1 + a^2 + a^4 + \ldots + a^{n-2}) (a^{n-1} + a^n) \tag{33}$$

Combining (31) and (33)

$$f[(n + 1)-2]=a^3 (1 + a + a^2 +\ldots+ a^{n-2}) (1 + a^2 + a^4 +\ldots+ a^{n-2})$$
$$+ a^3 (1 + a^2 + a^4 + \ldots + a^{n-2}) (a^{n-1} + a^n)$$
$$= a^3 (1 + a + a^2 + \ldots + a^{n-2} + a^{n-1} + a^n) (1 + a^2$$
$$+ a^4 + \ldots + a^{n-2}) \tag{34}$$

We substitute $(n + 1)$ in (14) where $(n + 1)$ is odd, since n is even by assumption. The powers are calculated first

$$[2\,(n+1)-3-(-1)^{n+1}]\,/\,2 = n \tag{35}$$

$$[2\,(n+1)-5-(-1)^{n}]\,/\,2 = n\text{-}2 \tag{36}$$

We have,

$$f\,[(n+1)\text{-}2] = a^3\,(1+a+a^2+\ldots+a^n)\,(1+a^2+a^4+\ldots+a^{n-2}) \tag{37}$$

But (37) is identically the same as (34). This completes the proof of Theorem 1.

The combinatorial generating functions are needed to calculate the percentage points for the slippage tests. As a simple demonstration consider the case of drawing two ranks at random from ranks 1, 2, ..., 6. The possible combinations and their sums are given by (14).

$$f\,(6\text{-}2) = a^3\,(1 + a + a^2 + a^3 + a^4)\,(1 + a^2 + a^4) \tag{38}$$

or in a slightly different form,

$$f\,(6\text{-}2) = a^3\,\left(\sum_{y=0}^{4} a^y\right)\,\left(\sum_{z=0}^{2} a^{2z}\right) \tag{39}$$

In this product of two series each term has a coefficient of 1. For a rank sum of T_j we can write,

$$T_j = 3 + y + 2z \tag{40}$$

$$y = T_j - 3 - 2z \tag{41}$$

The series in (39) have terms only for

$$y \leq 4 \;,\; z \leq 2 \tag{42}$$

$$y, z \geq 0 \quad \text{and} \quad y, z \text{ integers} \tag{43}$$

The value of y in (41) can be substituted in (39) thereby resulting an expression only in the variable z. But it seems better to conduct a search as shown below.

Thus, using (38), the number of ways of getting a $T_j = 8$ can be calculated. We start with $z = 0$ and from (41) we get, $y = 8\text{-}3 = 5$. But y=5, according to (42) it is not permissible. Hence it is not possible to form a rank sum of 8 by taking $z = 0$. Next we take $z = 1$, from (41), we obtain $y = 8\text{-}3\text{-}2 = 3$, it is permissible. Thus we have obtained one combination to give a rank sum of 8. In the third step we take $z = 2$, and using it in (41), we obtain $y = 8\text{-}3\text{-}4 = 1$. This value of y is also permissible. Thus we have obtained another combination to

to give a rank sum of 8, next we take $z = 3$ and from (41) $y = -1$ which violates (43), and we stop here.

Therefore, we arrive at the conclusion that the probability of getting a rank sum of 8, in a random assignment of two ranks drawn from possible choices of 1, 2, ..., 6 is $2 / C^6_2 = 2 / 15$.

In a problem where the number of combinations is large, it would be practically impossible to enumerate all possible rank sums and then calculate the probability of getting a particular one. Thus, a feasible solution to the general problem seems to be along the lines shown above.

If the possible choices of ranks are extended to include 1, 2, ..., n, taken two at time, the combinatorial generating function takes the form.

$$f(n-2) = a^3 \left(\sum_{y=0}^{l} a^y \right) \left(\sum_{z=0}^{m} a^{2z} \right) \tag{44}$$

where $l = [2n-3-(-1)^n]/2$, $m = [2n-5-(-1)^{n-1}]/4$ (45)

From (44) we infer that the maximum possible rank sum of two ranks is given by

$$T_{max} = 3 + [2n-3-(-1)^n]/2 + [2n-5-(-1)^{n-1}]/2 = 2n-1$$

Thus $T_j \leq 2n-1$. For y and z we can also impose the following restrictions

$$y, z \geq 0 \tag{46}$$

$$y \leq [2n-3-(-1)^n]/2 \tag{47}$$

$$z \leq [2n-5-(-1)^{n-1}]/4 \tag{48}$$

In this problem, if it is required to find the possible ways of forming a particular tank sum T_0, we again start with $z = 0$ and obtain $y = T_0 -3-2z$. If it is permissible with respect to (46) and (47), then it means that we have obtained a combination to give a rank sum of T_0. At the next step we take $z = 1$ and continue until all the possible ways of getting a rank sum of T_0 are exhausted. Note that this method of search usually involves only one tail of the symmetric distributions of the rank sums, and can be carried out in a reasonable amount of time as the applications show.

We continue our examples with the case where there are replications. Supposing the experiment of our prevoius exəmple, is rep-

licated R times. Thus the combinatiorial generating function will rake the form,

$$f(n-2, R) = \left[a^3 (1 + a + a^2 + \ldots + a^{\frac{2n-3-(-1)^n}{2}}). \right.$$

$$\left. (1 + a^2 + a^4 + \ldots + a^{\frac{2n-5-(-1)^{n-1}}{2}}) \right]^R \qquad (49)$$

In this form $f(n-2, R)$ is not suitable for our iterative procedure for determining the number of combinations for a particular rank sum. We therefore put it in a different form to meet our requirements.

$$f(n-2, R) = a^{3R} \left(\frac{1-a^p}{1-a} \right)^R \left(\frac{1-a^w}{1-a^2} \right)^R \qquad (50)$$

where $p = [2n-1-(-1)^n]/2$ and $w = [2n-1-(-1)^{n-1}/2$ \qquad (51)

We can also express (50) using the binomial series expansions given below:

$$(1-a^p)^R = \sum_{x=0}^{\infty} (-1)^x \binom{R}{x} a^{px} \qquad (52)$$

$$(1-a^w)^R = \sum_{y=0}^{\infty} (-1)^y \binom{R}{y} a^{wy} \qquad (53)$$

$$(1-a)^{-R} = \sum_{z=0}^{\infty} \binom{R+z-1}{z} a^z \qquad (54)$$

$$(1-a^2)^{-R} = \sum_{v=0}^{\infty} \binom{R+v-1}{v} a^{2v} \qquad (55)$$

The series in (52) and (53) have finite number of terms for R a positive integer, and R is a positive integer in our case. The series in (54) and (55) are convergent for $a^2 < 1$. But this is quite immaterial in our case. Because we are only interested in the powers of a's resulting from the product of these series, and we pay no attention whatsoever to the series being convergent or divergent.

Now a search for a particular rank sum T_j will be conducted in (50), where series expansions shown in (52), (53), (54) and (55) have been substituted. Note that there are four variables x, y. z and v and

they all have to be nonnegative integers, and we must also have $x, y \leq R$.

For a given T_j we have

$$T_j = 3R + (x/2)[2n\text{--}1\text{--}(\text{--}1)^n] + (y/2)[2n\text{--}1\text{--}(\text{--}1)^{n-1}] + z + 2v \quad (56)$$

The sum of the coefficients of terms containing a particular (a^{T_j}) will be calculated from (50).

For finding the number of ways of getting a particular rank sum T_j, a search for all permissible values of x, y, z and v must be conducted. This procedure will be demonstrated by examples in the subsequent discussions. We now state another theorem that might be useful in determining the combinatorial generating functions.

Theorem 2. The rank sums for all possible combinations of ranks $1, 2, 3, \ldots, n$, taken three at a time, are given by the following combinatorial generating function:

$$f(n\text{--}3) = \sum_{i=3}^{n} a^{i+3}\left(1 + a + a^2 + \ldots + a^{\frac{2i-5+(-1)^i}{2}}\right).$$

$$\left(1 + a^2 + a^4 + a^6 + \ldots a^{\frac{2i-7+(-1)^{i-1}}{2}}\right) \quad (58)$$

where $i = 1, 2, \ldots$ positive integers. We will prove (58) by induction.

1) We will assume (58) to be true for n and then prove that it will also be true for $(n + 1)$.

For n, (58) takes the following form:

$$f(n\text{--}3) = a^6 + a^7(1 + a + a^2) + a^8(1 + a + a^2)(1 + a^2) + \ldots$$

$$\ldots + a^{n+3}\left(1 + a + a^2 \ldots + a^{\frac{2n-5+(-1)^n}{2}}\right)$$

$$\left(1 + a^2 + a^4 + \ldots + a^{\frac{2n-7+(-1)^{n-1}}{2}}\right) \quad (59)$$

If we now increase n to $(n + 1)$ and form combinations of three, these combinations will contain all the combinations of three for n, plus the combinations of the rank $(n + 1)$ with the combinations of

two taken from ranks 1, 2, ..., n. Hence the additional combinations of three ranks caused by gioing from n to (n + 1) are equal to

a^{n+1} f (n–2) or using (14) for f (n–2),

$$a^{n+1} \text{ f } (n-2) = a^{(n+1)+3} \left(1 + a + a^2 + \ldots + a^{\dfrac{2(n+1)-5+(-1)^{n+1}}{2}}\right).$$

$$\left(1 + a^2 + a^4 + \ldots + \ldots + a^{\dfrac{2(n+1)-7+(-1)^{(n+1)-1}}{2}}\right) \quad (60)$$

where we have used, $-(-1)^n = (-1)^{n+1}$ and $-(-1)^{n-1} = (-1)^{(n+1)-1}$

Now summing (59) and (60), we get the expression for f [(n + 1)–3]

f [(n + 1) –3] = $a^6 + a^7 (1 + a + a^2) + a^8 (1 + a + a^2) (1 + a^2) + \ldots$

$$+ a^{n+3} \left((1 + a + a^2 + \ldots + a^{\dfrac{2n-5+(-1)^n}{2}}\right).$$

$$\left(1 + a^2 + a^4 + \ldots + a^{\dfrac{2n-7+(-1)^{n-1}}{2}}\right) + a^{(n+1)+3} \left(1 + a\right.$$

$$+ a^2 + \ldots + a^{\dfrac{2(n+1)-5+(-1)^{n+1}}{2}}\right) \left(1 + a^2\right.$$

$$+ a^4 + \ldots + a^{\dfrac{2(n+1)-7+(-1)^{(n+1)-1}}{2}}\right) \quad (61)$$

But (61) is exactly what we would get if we had used (n + 1) in (58). To complete the proof of Theorem 2,

2) We have to prove that (58) is ture for the smallest meaningful value of n, which is 3 in this case. We know that for n = 3, that is for ranks 1, 2, 3 the rank sum of three ranks will be 6, and this is the only sum of three ranks we can get in this case. Applying (58)

$$\text{f } (3-3) = \sum_{i=3}^{3} \quad a^{3+3} (1) (1) = a^6$$

This completes the proof of Theorem 2. Replications are treated the same way for f (n–3) as was done for f (n–2). For R replications of the experiment the combinatorial generating function is given by the R th power of f (n–3).

$$f(n-3, R) = [f(n-3)]^R \tag{62}$$

Theorem 1 and 2 would be useful in determining the combinatorial generating function for n ranks taken two or three at a time. However we also need a more general theorem for obtaining the combinatorial generating function for any given situation.

Theorem 3. The combinatorial generating function of n ranks taken k at a time is given by the following recurrence relation

$$f(n-k) = \sum_{i=k-1}^{n-1} a^{i+1} f[i-(k-1)] \tag{63}$$

We will prove Theorem 3 also by induction.

1) Assuming (63) to be true for n we will show that it is also true for (n + 1). $f[(n+1)-k]$ is the sum of f (n–k) and the new additional combinations caused by increasing n to (n + 1). But the new combinations will be formed by the combinations of ranks 1, 2, 3, ..., n, taken (k–1) at a time and the rank (n + 1). That is

$$f[(n+1)-k] = \sum_{i=k-1}^{n-1} a^{i+1} f[i-(k-1)] + a^{n+1} f[n-(k-1)] \tag{64}$$

The first term with the summation sign on the right side of (64) is assumed valid by assumption and is given by (63). The second term shows the additional combinations caused by increasing n to (n + 1). But the second term can also be taken under the summation sign by increasing the upper limit of summation from (n–1) to n. Hence,

$$f[(n+1)-k] = \sum_{i=k-1}^{n} a^{i+1} f[i-(k-1)] \tag{65}$$

But (65) is precisely the same expression that would result if (n + 1) were used in (63).

2) Next we have to prove that (63) is true for the smallest meaningful value of n = k.

From (63),

$$f(k-k) = \sum_{i=k-1}^{k-1} a^{i+1} f[i-(k-1)] = a^k f[(k-1)-(k-1)] \tag{66}$$

But if we take (k–1) of the 1, 2, ..., (k–1) ranks, the rank sum will be,

$$f\left[(k-1)-(k-1)\right] \;=\; a^{k\ (k-1)/2} \tag{67}$$

and from (66)

$$f(k-k) \;=\; a^k \cdot a^{k\ (k-1)/2} = a^{k\ (k+1)/2} \tag{68}$$

But the sum of k ranks 1, 2, ..., k is $k(k+1)/2$ and its symbol in our notation is,

$$a^{k\ (k+1)/2} \tag{69}$$

(68) and (69) are identical. The proof of Theorem 3 is completed.

CONCLUSIONS

Note that by the use of Theorems 1, 2 and 3 the combinatorial generating functions for any two way layout problem might be obtained. In this process, Theorems 1 or 2 might be useful in preparing the starting combinatorial generating functions for the application of the recurrence relation given in Theorem 3.

In any significance testing, the interest is usually in the tail areas of the distribution of rank sums. Thus, instead of studying all the terms of the generating function for a large problem, it would be sufficient just to look at the tail probabilities. For this reason it would be better to put the combinatorial generating functions in a special form. However, due to shortage of space, this topic will be discussed by the author in a different paper in the future.

It should also be noted here that the application of Theorems 1, 2, and 3 are by no means limited to the rank sum problems in a two way layout. In fact, they might be useful in any area where combinatorial problems of similar nature might arise.

REFERENCES

ACAR, B.S. and PETTITT, A.N. 1985. Perdiction Using Nonparametric Techniques: Some Comparisons, *J. Statist. Comput. Simul.* Vo. 21, pp. 117–134.

CAMPBELL, G., and SKILLINGS, J.H. 1985. Nonparametric stepwise Multiple Comparison Procedures, *J. Amer. Statist. Assoc.* Vol. 80, 998–1003.

CUZICK, J. 1985. Asymptotic Proprties of Censored Linear Rank Tests, *The Annals of Statistics*, Vol. 13, pp. 133–141.

FRIEDMAN, M. 1937. The Use of Ranks to Avoid the Assumption ogf Normality Implicit in the Analysis of Variance, *Journal American Statistical Association*, 32, pp. 675–699.

HENERY, R.J. 1986. Interpretation of Average Ranks, *Biometrika*, 73, pp. 224–227.

HODGES, J.D., and LEHMANN, E.L. 1962. Rank Methods for Combination of Independent Experiments in Analysis of Variance, *Ann. Math. Statist.* 33, pp. 482–497.

HORA, S.C., and CANOVER, W.J. 1984. The F-statistic in the two-way Layout with Rank-Score Transformed Data, *Jour. Amer. Statist. Assoc.* Vol. 79. pp. 668–673.

KENDALL, M.G. 1970. "Rank Correlation Methods", 4th Ed. Hafner, New York.

MAK, T.K. 1985. Interpretation and Estimation in Ranking Theory, *Biometrika*, 72, pp. 227–229.

MEHRA, K.L., and SEN, P.K. 1969. On a Class of Conditionally Distribution-Free Tests for Interactions in Factorial Experiments, *Annals of Mathematical Statistics*, 40, pp. 658–664.

ÖZKAN, F., 1975. "Sıra Toplamları Kullanarak Hipotez Testleri". An unpublished Thesis, *Middle* East Technical University, Ankara, Turkey.

SALTER, K.C. and FAWCETT, R.F. 1985. A Robust and Powerful Rank Test of Treatment Effects in Balanced Incomplete Block Desings, *Commun. Statist,-Simyula.* Computa, 14 (4), pp. 807–828.

SCHLUCHTER, M.D. 1985. An Aligned Rank Test for Censored Data from Randomized Block Designs, *Biometrika*, 72, pp. 609–618.