Statistica Sinica 3(1993), 277-294

ON THE NUMBER OF RUNS AND RELATED STATISTICS

Markos V. Koutras and Stavros G. Papastavridis

University of Athens, Greece

Abstract: The enumeration of combinations with prescribed number of successions of specified length and a related occupancy problem are considered. The generating functions for both situations are derived and the relation between the linear and circular arrangements is obtained. The results are applied to a variety of well known problems, including Kaplansky's "Problème des Ménages", the study of Fibonacci and Lucas numbers and the reliability of consecutive k-out-of-n systems.

Key words and phrases: Run, Fibonacci numbers, Lucas numbers, problème des ménages, consecutive k-out-of-n systems.

1. Introduction

Many articles in the combinatorial literature consider the problem of the enumeration of combinations of the first n positive integers $\{1, 2, \ldots, n\}$, with certain restrictions on the appearances of successions of consecutive numbers. Kaplansky (1943), based the solution of the famous "Problème des Ménages", on the computation (by recurrence) of the number of combinations of n objects taken k at a time with no two selected objects being consecutive. Many years later Moser and Abramson (1979) considered combinations with restricted differences and cospan and gave explicit expressions for a large class of problems, including the abovementioned situation as a special case. A problem of the same kind, originating from engineering (short-circuiting of adjacent electrodes) was treated by Apostol (1988). Derman, Lieberman and Ross (1982), solved the problem of enumerating combinations with no k consecutive elements by considering an equivalent formulation through distributions of balls into urns. Recently Hwang and Yao (1991), considered a generalization of Kaplansky's result. Closely related are the problems studied by Hwang (1981) and Konvalina (1981).

In the situations mentioned above there are two possible arrangements of the n elements (integers): linear and circular (where, in the second case, 1 is considered to be the next integer following n). The observation which gave rise to the present paper was that in almost all the known cases there is a simple relation between the number of combinations in the linear and circular situations (see

Hwang and Yao (1991), Kaplansky (1943) and Riordan (1958)). Furthermore, the fact that the recurrence relations satisfied by certain quantities (for example, reliabilities of consecutive k-out-of-n systems) which are expressed directly via these numbers, are the same for both the linear and circular cases (see Du and Hwang (1988), Lambiris and Papastavridis (1985)), was considered a strong indication that some interesting mathematical relation is hidden between the two arrangements under very general assumptions on the restrictions imposed on the combinations to be enumerated.

With these facts in mind, we introduce (Section 2) the notion of run (linear and cyclic) within a combination of the first n positive integers, and treat the general problem of enumerating the combinations with specific number of prescribed runs. In the same Section, a simple trick is used to transform the initial problem into an equivalent occupancy model. In Section 3 we give the generating functions of the numbers appearing in the occupancy setting, while in Section 4 the results are converted to the original problem. Finally, in Section 5 we show how the outcomes of the previous Sections can be used either to derive directly a number of well known results or to treat some interesting generalizations of certain problems. Related results, based on a somewhat different definition of runs can be found in Fu (1993), Godbole (1990), Hirano (1986), Ling (1989), Papastavridis (1990) and Philippou and Makri (1986). Finally, we mention that these results have clear potential of statistical applications along the lines suggested by Agin and Godbole (1991), something that we plan to deal with in a subsequent paper.

2. Definitions and Notations

Let

$$1 \leq x_1 < x_2 < \cdots < x_r \leq n \tag{1}$$

be an r-combination of the n consecutive integers $\{1, 2, \ldots, n\}$ and a a positive integer. The a-combination

$$x_j < x_{j+1} < \cdots < x_{j+a-1}$$

will be called a *run of length* a (of the original r-combination), if one of the following situations arises

a. $x_{j+1} = x_j + 1$, $j = 1, 2, \ldots, a-1$ and $x_{a+1} \neq x_a + 1$

b. $x_{j+1} = x_j + 1$, $j = i, i+1, \ldots, i+a-2$ and $x_{j+1} \neq x_j + 1$, $j = i-1, i+a-1$ $(i > 1)$

c. $x_{j+1} = x_j + 1$, $j = r-a+1, \ldots, r-1$ and $x_{r-a+1} \neq x_{r-a} + 1$.

Also let A be any set of positive integers. A run whose length a belongs to A will be called an A-run. In the case where the n consecutive integers $\{1, 2, \ldots, n\}$ are considered cyclicly arranged (i.e. 1 is viewed as the next integer following n) we may similarly introduce the terms of *cyclic run of length* a and *cyclic* A*-run (the run* is defined now by condition (b) alone, assuming that the addition is performed mod n).

For $n \geq r > 0$ let $N(n, r, s)$ denote the number of the r-combinations of the integers $\{1, 2, \ldots, n\}$ which contain exactly s A-runs (s is any non negative integer) in the linear case. Assuming, conventionally, that

$$N(n, 0, s) = \delta_{s0} = \begin{cases} 1 & \text{if} \quad s = 0, \\ 0 & \text{if} \quad s > 0, \end{cases}$$

we may express the total number of combinations with exactly s A-runs in the form

$$N(n, s) = \sum_{r=0}^{n} N(n, r, s), \quad n > 0.$$

Finally, in order to have the aforementioned quantities defined for any nonnegative n, r and s, set

$$N(0, s) = N(0, 0, s) = \delta_{s0}.$$

For the circular case, denote by $N_c(n, r, s)$ the corresponding numbers. Introducing the convention

$$N_c(n, 0, s) = \delta_{s0} \quad n > 1,$$

we may express the total number of (cyclic) combinations with exactly s A-runs in the form

$$N_c(n, s) = \sum_{r=0}^{n} N_c(n, r, s), \quad n > 1.$$

We complete the range of values of n and r by assuming that

$$N_c(1, s) = N_c(1, 0, s) = \delta_{s0}, \quad N_c(0, s) = N_c(0, 0, s) = 0, \quad s \geq 0.$$

In this paper we shall present a unified study of these sequences of numbers, emphasizing on the relations between the linear and cyclic cases. It should be mentioned that many of well known problems can be viewed as special cases of the general setting given above. For example:

a. If $A = \{2, 3, \ldots\}$, the numbers $N(n, r, 0)$, $N_c(n, r, 0)$ are the quantities used by Kaplansky (1943) for the solution of the "Problème des Ménages", while $N(n, 0)$, $N_c(n, 0)$ are the Fibonacci and Lucas numbers respectively. The more general numbers $N(n, r, s)$ were studied by Riordan (1968, page 11).

b. If $A = \{k, k+1, \ldots\}$, the numbers $N(n, 0)$, $N_c(n, 0)$ coincide with the Fibonacci and Lucas numbers of order k respectively (Charalambides (1991), Hwang and Yao (1991)). In this case the sums

$$R(p,n) = \sum_{r=0}^{n} N(n,r,0) q^r p^{n-r}$$
$$R_c(p,n) = \sum_{r=0}^{n} N_c(n,r,0) q^r p^{n-r} \qquad n \geq k, \quad p = 1-q$$

give the reliability of a consecutive k-out-of-n system with component failure probabilities q (Derman, Lieberman and Ross (1982)).

c. If $A = \{k\}$, then $N(n, s)$ gives the number of binary vectors in n-space with exactly s isolated k-tuples, which was studied by Apostol (1988).

d. The study of the so called strict consecutive-k-out-of-n system (Bollinger (1985)) corresponds to the case $A = \{1, 2, \ldots, k-1\}$.

In order to study the generating functions of the numbers $N(n, r, s)$, $N_c(n, r, s)$ it is convenient to transform our models to some equivalent occupancy models. For this purpose, to every r-combination (1) assign a place-indicator vector $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ in n-space defined by

$$\varepsilon_i = \begin{cases} 0 & \text{if} \quad i = x_1, x_2, \ldots, x_r, \\ 1 & \text{otherwise.} \end{cases}$$

In the linear case, the $n-r$ 1's determine $n-r+1$ cells to which the r 0's are distributed. It is obvious that from every A-run of the r-combination, a cell is created containing $a \in A$ objects (0's). In the sequel, such a cell will be called A-cell. Denoting by $M(r, m, s)$ the number of ways of distributing r like objects into m different cells with exactly $s \leq m$ of the cells being A-cells, we obtain

$$N(n,r,s) = M(r, n-r+1, s) \quad 0 < r \leq n, \tag{2a}$$
$$M(r,m,s) = N(m+r-1, r, s) \quad m > 0, \; r > 0. \tag{2b}$$

In the cyclic case, we have again

$$N_c(n,r,s) = M_c(r, n-r+1, s) \quad 0 < r < n, \tag{3a}$$
$$M_c(r,m,s) = N_c(m+r-1, r, s) \quad m > 1, \; r > 0, \tag{3b}$$

with $M_c(r, m, s)$ denoting now the number of ways of distributing r like objects into m different cells with either exactly $s-1$ of the cells $2, 3, \ldots, m-1$ being A-cells and cells 1 and m containing together a *total* number of $a \in A$ objects or

s of the cells $2, 3, \ldots, m-1$ being A-cells and cells 1 and m containing together a *total* number of $a \notin A$ objects.

In order to extend the range of validity of (2a) and (2b), we introduce the following conventions for $M(r, m, s)$

$$\begin{aligned} M(0, m, s) &= \delta_{s0} \quad m \geq 1 \\ M(r, 0, s) &= 0 \qquad r \geq 1. \end{aligned} \tag{4}$$

Relations (2a) are now true for $0 \leq r \leq n$ and (2b) for $m \geq 0$, $r \geq 0$, $m + r > 0$. In the same way write

$$M_c(0, m, s) = \delta_{s0} \quad m \geq 2 \tag{5}$$

$$M_c(r, 1, s) = M_c(r, 0, s) = 0 \quad r \geq 0 \tag{6}$$

which allows the extention of (3a) to $0 \leq r < n$ and (3b) to $m \geq 0$, $r \geq 0$, $m + r > 0$. One more convention, used in the sequel with no reference to (2a) or (2b) is the following

$$M(0, 0, 0) = 1. \tag{7}$$

For the study of the numbers mentioned above we shall make use of the enumerating generating function (enumerator) of the A-cells, namely

$$f(t) = \sum_{i \in A} t^i$$

as well as its complement

$$g(t) = \sum_{i \notin A} t^i = \frac{1}{1-t} - f(t), \quad |t| < 1$$

which enumerates the non A-cells.

3. Generating Functions of the Numbers $M(r, m, s)$ and $M_c(r, m, s)$

In this section we derive the generating functions of the numbers $M(r, m, s)$ and $M_c(r, m, s)$ (with respect to several combinations of the parameters r, m, s), recurrence relations, and the connection between the linear and cyclic cases. The results are presented in the following propositions.

Lemma 1. *The generating function $G(t; m, s)$ of the numbers $M(r, m, s)$, $r = 0, 1, \ldots$ is given by*

$$G(t; m, s) = \sum_{r=0}^{\infty} M(r, m, s) t^r = \binom{m}{s} f^s(t) g^{m-s}(t) \quad m \geq s. \tag{8}$$

Proof. The enumerator for occupancy of a specific cell is $f(t)$ if it is an A-cell and $g(t)$ if it is not an A-cell. Since we require s of the m distinguishable cells to be A-cells and the rest $m-s$ non A-cells we easily deduce the expression (8).

Relation (8) justifies the use of initial condition (7) which, at first sight, seems unreasonable (it contradicts the translation equalities (2a) and (2b)). Also, because of (8) we may write

$$G(t;m,s) = \binom{m}{s} G(t;s,s)G(t;m-s,0),$$

and expanding the generating functions on both sides we obtain

$$\begin{aligned} M(r,m,s) &= \binom{m}{s} \sum_{i=0}^{r} M(i,s,s)M(r-i,m-s,0) \\ &= \binom{m}{s} \sum_{i=0}^{r} M(r-i,s,s)M(i,m-s,0). \end{aligned}$$

Theorem 1. *The double generating function $G(t,x;s)$ of the numbers $M(r,m,s)$, $r=0,1,\ldots$, $m=s,s+1,\ldots$ is given by*

$$G(t,x;s) = \sum_{m=s}^{\infty}\sum_{r=0}^{\infty} M(r,m,s)t^r x^m = [xf(t)]^s[1-xg(t)]^{-s-1}. \tag{9}$$

Proof. By making use of the result of Lemma 1,

$$G(t,x;s) = \sum_{m=s}^{\infty} G(t;m,s)x^m = [xf(t)]^s \sum_{m=s}^{\infty} \binom{m}{s} [xg(t)]^{m-s},$$

and the assertion of the theorem is an immediate consequence of the well known identify

$$\sum_{m=s}^{\infty} \binom{m}{s} w^{m-s} = \frac{1}{s!}\frac{d^s}{dw^s}(1-w)^{-1} = (1-w)^{-s-1}. \tag{10}$$

The generating function of the above theorem could have been deduced along lines parallel to Goulden and Jackson (1983) (see Section 2.4). Note also the following relations which are direct consequences of formula (9)

$$G(t,x;s) = [xf(t)]^s G^{s+1}(t,x;0), \quad G(t,x;0) = [1-xg(t)]^{-1},$$

$$G(t,x;s) = \frac{xf(t)}{1-xg(t)} G(t,x;s-1), \quad s>0.$$

Corollary 1. a. *The triple generating function of the numbers $M(r,m,s)$ is given by*

$$\sum_{s=0}^{\infty}\sum_{m=s}^{\infty}\sum_{r=0}^{\infty} M(r,m,s)t^r x^m z^s = \frac{1}{1-x[g(t)+zf(t)]}. \tag{11}$$

b. *The double generating function of the numbers $M(r,m,s)$, $r=0,1,\ldots$, $s=0,1,\ldots$ is given by*

$$\sum_{s=0}^{m}\sum_{r=0}^{\infty} M(r,m,s)t^r z^s = [g(t)+zf(t)]^m. \tag{12}$$

Proof. a. It suffices to multiply (9) by z^s, sum up for $s=0,1,\ldots$ and make use of the geometric series expansion.
b. The result is easily derived either by expanding the right hand side of (11) in a power series with respect to x or by multiplying (8) by z^s, summing up for $s=0,1,\ldots$ and making use of the binomial formula.

The next Theorem is of great importance, since it provides the main link between the linear and cyclic cases.

Theorem 2. *The generating function $G_c(t;m,s)$ of the numbers $M_c(r,m,s)$, $r=0,1,\ldots$ is given by*

$$\begin{aligned} G_c(t;m,s) &= \sum_{r=0}^{\infty} M_c(r,m,s)t^r \\ &= \frac{t}{m-1}G'(t;m-1,s)+G(t;m-1,s) \quad m>\max(1,s), \quad (13) \\ G_c(t;1,0) &= 0 \end{aligned}$$

(G' denotes here the derivative of G with respect to t).

Proof. Assume first that $m-1>s\geq 1$. In this case the cyclic distributions of interest (i.e. the ones that are enumerated by $G_c(t;m,s)$) may be partitioned into two disjoint classes according to whether the total number a of objects contained in cells 1 and m belongs to A or not. If a total number of i balls are used for the first and mth cells, it is evident that there exist $i+1$ possible assignments to those two cells. Therefore, the joint enumerator of the first and mth cell is

$$\sum_{i\notin A}(i+1)t^i = tg'(t)+g(t)$$

if $a\notin A$, and

$$\sum_{i\in A}(i+1)t^i = tf'(t)+f(t)$$

if $a \in A$. Also the enumerator of the other $m-2$ cells is given by $G(t;m-2,s)$ if $a \notin A$, and $G(t;m-2,s-1)$ if $a \in A$. Combining all these arguments we obtain

$$\begin{aligned} G_c(t;m,s) &= \binom{m-2}{s} f^s(t) g^{m-s-2}(t)[tg'(t)+g(t)] \\ &\quad + \binom{m-2}{s-1} f^{s-1} g^{m-s-1}(t)[tf'(t)+f(t)] \end{aligned}$$

which, after some algebra, yields

$$\begin{aligned} G_c(t;m,s) &= \left\{ \binom{m-2}{s} + \binom{m-2}{s-1} \right\} f^s(t) g^{m-s-1}(t) \\ &\quad + \frac{t}{m-1} \binom{m-1}{s} \left\{ f^s(t) \frac{d}{dt} g^{m-s-1}(t) + g^{m-s-1}(t) \frac{d}{dt} f^s(t) \right\} \\ &= G(t;m-1,s) + \frac{t}{m-1} G'(t;m-1,s). \end{aligned}$$

If $1 \le s = m-1$ or $0 = s < m-1$ it is easy to verify that

$$G_c(t;m,s) = f^{m-2}(t)(tf'(t)+f(t))$$

and

$$G_c(t;m,s) = g^{m-2}(t)(tg'(t)+g(t))$$

respectively, which agree with (13). Finally, for $0 = s = m-1$ the initial condition $G_c(t;1,0) = 0$ is immediately derived from (6).

Relation (13) implies that the generating functions $G_c(t;m,s)$ and $G(t;m-1,s)$ have the same constant term. This fact explains assumption (5). Note also that, by differentiating formula (8) with respect to t, we may prove that

$$G'(t;m,s) = mG(t;m-1,s-1)f'(t) + mG(t;m-1,s)g'(t),$$

and substituting $G'(t;m-1,s)$ in (13) we obtain the next alternative expression for the generating function $G_c(t;m,s)$

$$\begin{aligned} G_c(t;m,s) &= tf'(t)G(t;m-2,s-1) + tg'(t)G(t;m-2,s) \\ &\quad + G(t;m-1,s) \qquad 0 < s < m-1. \end{aligned} \tag{14}$$

It should be mentioned that, the cyclic case studied here, has a certain similarity with the "logarithmic connection" discussed in Goulden and Jackson (1984).

Theorem 3. *The double generating function $G_c(t,x;s)$ of the numbers $M_c(r,m,s)$, $r=0,1,\dots$, $m=s+1,s+2,\dots$ is given by*

$$G_c(t,x;s) = x^2tf'(t)G(t,x;s-1) + x(txg'(t)+1)G(t,x;s) \quad s \geq 1, \tag{15}$$

$$G_c(t,x;0) = \frac{x^2(tg'(t)+g(t))}{1-xg(t)}. \tag{16}$$

Proof. We have

$$\begin{aligned} G_c(t,x;s) &= \sum_{m=s+1}^{\infty}\sum_{r=0}^{\infty} M_c(r,m,s)t^r x^m = \sum_{m=s+1}^{\infty} G_c(t;m,s)x^m \\ &= G_c(t;s+1,s)x^{s+1} + \sum_{m=s+2}^{\infty} G_c(t;m,s)x^m. \end{aligned}$$

For $s \geq 1$ make use of the expression (see (8) and (13))

$$G_c(t;s+1,s) = \frac{t}{s}G'(t;s,s) + G(t,s,s) = (tf'(t)+f(t))f^{s-1}(t)$$

and (14), and after some rather simple algebraic manipulations (15) is obtained. For $s=0$ it suffices to employ formulas

$$G_c(t;1,0) = 0, \quad G_c(t;m,0) = (tg'(t)+g(t))g^{m-2}(t).$$

Relations (9) and (15) could be combined to obtain a direct expression for $G_c(t,x;s)$. Specifically,

$$G_c(t,x;s) = \frac{x^{s+1}f^{s-1}(t)}{[1-xg(t)]^{s+1}}\left\{tf'(t)+f(t)+tx(g'(t)f(t)-f'(t)g(t))\right\} \quad s>0. \tag{17}$$

Corollary 2. a. *The triple generating function of the numbers $M_c(r,m,s)$ is given by*

$$\sum_{s=0}^{\infty}\sum_{m=s+1}^{\infty}\sum_{r=0}^{\infty} M_c(r,m,s)t^r x^m z^s = \frac{x^2[tg(t)+tzf(t)]'}{1-x[g(t)+zf(t)]}. \tag{18}$$

b. *The double generating function of the numbers $M_c(r,m,s)$, $r=0,1,\dots$, $s=0,1,\dots,m-1$ is given by*

$$\sum_{s=0}^{m-1}\sum_{r=0}^{\infty} M_c(r,m,s)t^r z^s = [tg(t)+tzf(t)]'[g(t)+zf(t)]^{m-2}, \quad m>1 \tag{19}$$

(*The derivative here is taken with respect to t*).

Proof. Expression (18) is easily derived if we multiply (14) or (17) by z^s, sum for $s = 1, 2, \ldots$ and append the term $G_c(t, x; 0)z^0$ from (16). For (19), expand (18) in a power series with respect to x and consider the coefficient of x^m.

4. Generating Functions of the Numbers $N(n, r, s)$ and $N_c(n, r, s)$

The generating functions of the numbers $N(n, r, s)$ and $N_c(n, r, s)$ are directly expressed in terms of the corresponding generating functions of the numbers $M(r, m, s)$ and $M_c(r, m, s)$ as the next propositions show.

Theorem 4. *The double generating function $F(t, x; s)$ of the numbers $N(n, r, s)$ with respect to r and n is given by*

$$F(t, x; s) = \sum_{n=s}^{\infty} \sum_{r=0}^{n} N(n, r, s) t^r x^n = \frac{1}{x} G(tx, x; s), \quad s \geq 1 \tag{20}$$

$$F(t, x; 0) = \sum_{n=0}^{\infty} \sum_{r=0}^{n} N(n, r, 0) t^r x^n = \frac{1}{x} \{G(tx, x; 0) - 1\}. \tag{21}$$

Proof. For $s \geq 1$ we have, in view of (4),

$$G(t, x; s) = \sum_{r=0}^{\infty} \left\{ \sum_{m=s}^{\infty} M(r, m, s) x^m \right\} t^r = \sum_{r=1}^{\infty} \left\{ \sum_{m=s}^{\infty} M(r, m, s) x^m \right\} t^r,$$

and employing the transformation $r + m - 1 = n$ in the inner sum, we obtain the expression

$$G(t, x; s) = \sum_{r=1}^{\infty} \sum_{n=r+s-1}^{\infty} M(r, n - r + 1, s) x^{n-r+1} t^r$$

which, by interchanging the order of summation, and making use of (2a), takes the form

$$G(t, x; s) = x \sum_{n=s}^{\infty} \sum_{r=1}^{n-s+1} N(n, r, s) t^r x^{n-r}.$$

Since

$$N(n, 0, s) = 0, \quad N(n, r, s) = 0 \qquad r = n - s + 2, \ldots, n$$

we deduce that

$$G(tx, x; s) = x \sum_{n=s}^{\infty} \sum_{r=0}^{n} N(n, r, s) t^r x^n = x F(t, x; s)$$

and the proof of (20) is over. For the proof of (21) we may start from

$$G(t, x; 0) = \sum_{r=0}^{\infty} \left\{ \sum_{m=0}^{\infty} M(r, m, 0) x^m \right\} t^r = \sum_{r=0}^{\infty} \left\{ M(r, 0, 0) + \sum_{m=1}^{\infty} M(r, m, 0) x^m \right\} t^r$$

which, in view of (4) and (7), becomes

$$G(t,x;0) = 1 + \sum_{r=0}^{\infty}\left\{\sum_{m=1}^{\infty} M(r,m,0)x^m\right\}t^r.$$

Employing analogous arguments with the ones used in the first part of the proof we may easily verify the truth of (21).

Combining Theorems 1 and 4 we deduce the following direct expression for the generating function $F(t,x;s)$

$$F(t,x;s) = \frac{x^{s-1}f^s(tx)}{[1-xg(tx)]^{s+1}} \qquad s \geq 1, \qquad F(t,x;0) = \frac{g(tx)}{1-xg(tx)}. \tag{22}$$

Corollary 3. *The generating function of the numbers*

$$N(n,s) = \sum_{r=0}^{n} N(n,r,s)$$

is given by

$$\begin{aligned} \sum_{n=s}^{\infty} N(n,s)x^n &= F(1,x;s) = \frac{1}{x}G(x,x,s) = \frac{x^{s-1}f^s(x)}{[1-xg(x)]^{s+1}} \qquad s>0, \\ \sum_{n=0}^{\infty} N(n,0)x^n &= F(1,x;0) = \frac{1}{x}\{G(x,x,0)-1\} = \frac{g(x)}{1-xg(x)}. \end{aligned} \tag{23}$$

Proof. It is immediate.

Theorem 5. *The double generating function $F_c(t,x;s)$ of the numbers $N_c(n,r,s)$ with respect to r and n is given by*

$$F_c(t,x;s) = \sum_{n=s+1}^{\infty}\sum_{r=0}^{n-1} N_c(n,r,s)t^r x^n = \frac{1}{x}G_c(tx,x;s) \qquad s \geq 0. \tag{24}$$

Proof. Similar to the proof of Theorem 4.

Corollary 4. *The generating function of the numbers*

$$N_c(n,s) = \sum_{r=0}^{n-1} N_c(n,r,s) \qquad n > s$$

is given by

$$\sum_{n=s+1}^{\infty} N_c(n,s)x^n = F_c(1,x;s) = \frac{1}{x}G_c(x,x;s) \qquad s \geq 0. \tag{25}$$

Proof. It is immediate.

5. Applications

In this section we examine some interesting special cases which are related with well known combinatorial problems.

a. Connection between the number of certain restricted combinations in the linear and circular cases

Substituting the power series expansions

$$G(t;m,s)=\sum_{r=0}^{\infty}M(r,m,s)t^r,\quad G_c(t;m,s)=\sum_{r=0}^{\infty}M_c(r,m,s)t^r$$

in identity (13) we deduce that

$$\begin{aligned}\sum_{r=0}^{\infty}M_c(r,m,s)t^r &= \frac{t}{m-1}\sum_{r=1}^{\infty}M(r,m-1,s)rt^{r-1}+\sum_{r=0}^{\infty}M(r,m-1,s)t^r\\ &= \sum_{r=0}^{\infty}\left\{\frac{r}{m-1}+1\right\}M(r,m-1,s)t^r\end{aligned}$$

which proves the following simple relation between the linear and circular distribution of r like objects into m different cells (with the conditions given in paragraph 2)

$$M_c(r,m,s)=\frac{r+m-1}{m-1}M(r,m-1,s)\qquad 0\le r<m,\quad m\neq 1. \tag{26}$$

The corresponding relation between numbers $N(n,r,s)$ and $N_c(n,r,s)$, is given by

Theorem 6.

$$N_c(n,r,s)=\frac{n}{n-r}N(n-1,r,s)\qquad 0\le s\le n-r,\quad r\neq 1. \tag{27}$$

Proof. Immediate consequence of (2), (3) and (26).

Theorem 6 has been proved by Kaplansky (1943) in the special case $A=\{2,3,\ldots\}$, $s=0$ and by Hwang and Yao (1991) for $A=\{k,k+1,\ldots\}$, $s=0$ (k any positive integer).

b. Binomial moments of certain occupancy distributions

Consider the following occupancy problem: r like objects are distributed into m distinct cells, arranged in a line and let X (a random variable) be the number of A-cells i.e. X counts the number of cells containing $a \in A$ objects. The probability function of X is given by

$$p(s;r,m) = Pr[X = s] = \frac{M(r,m,s)}{\binom{r+m-1}{r}} \qquad s = 0,1,\dots,m$$

and the next Theorem supplies a formula for the corresponding binomial moments

$$B_i(r,m) = E\left[\binom{X}{i}\right] = \sum_{s=0}^{\infty}\binom{s}{i}p(s;r,m).$$

Theorem 7. *If $a_r(m,i)$ is the coefficient of t^r in the expansion of $f^i(t)(1-t)^{-m+i}$ in a Taylor series i.e.*

$$\frac{f^i(t)}{(1-t)^{m-i}} = \sum_{r=0}^{\infty} a_r(m,i)t^r \tag{28}$$

then the ith binomial moment of the random variable X is given by

$$B_i(r,m) = \frac{\binom{m}{i}a_r(m,i)}{\binom{r+m-1}{r}} \qquad i = 1,2,\dots \tag{29}$$

Proof. Replacing z by $z+1$ in (12) we get

$$\sum_{r=0}^{\infty}\sum_{s=0}^{m} M(r,m,s)t^r(1+z)^s = \left[g(t)+f(t)+zf(t)\right]^m = \left(\frac{1}{1-t}+zf(t)\right)^m$$

or equivalently

$$\sum_{r=0}^{\infty}\binom{r+m-1}{r}\sum_{s=0}^{m}p(s;r,m)t^r\sum_{i=0}^{s}\binom{s}{i}z^i = \sum_{i=0}^{m}\binom{m}{i}\frac{f^i(t)}{(1-t)^{m-i}}z^i.$$

Changing the order of summation in the left hand side and substituting (28) in the right hand side, it follows that

$$\sum_{i=0}^{m}\sum_{r=0}^{\infty}\binom{r+m-1}{r}\left\{\sum_{s=i}^{m}\binom{s}{i}p(s;r,m)\right\}t^r z^i = \sum_{i=0}^{m}\sum_{r=0}^{\infty}\binom{m}{i}a_r(m,i)t^r z^i,$$

which implies (29).

The special case $A = \{k\}$ (k non negative integer) was treated by Riordan (1958) (see problems 9 and 10, p.103). The result given there could be easily derived if we notice that

$$\frac{f^i(t)}{(1-t)^{m-i}} = \frac{t^{ik}}{(1-t)^{m-i}} = \sum_{r\geq ik}\binom{m+r-(k+1)i-1}{r-ik}t^r$$

and substitute

$$a_r(m,i) = \binom{m+r-(k+1)i-1}{r-ik}$$

in formula (29).

c. Enumeration of binary vectors in n-space containing prescribed subsets of consecutive ones

In this section we treat a generalization of the combinatorial problem proposed and solved by Apostol (1988). Let $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ be a binary vector in n-space, that is, each component ε_i is either 0 or 1. We say that the binary vector contains an *isolated a-tuple* of consecutive ones if for some index $i \in \{1, 2, \dots, n-a+1\}$ we have

$$\varepsilon_{i-1} = 0, \quad \varepsilon_i = \varepsilon_{i+1} = \cdots = \varepsilon_{i+a-1} = 1, \quad \varepsilon_{i+a} = 0$$

(in the cases $i = 1$ and $i = n - a + 1$ we ignore the first and last condition respectively).

Introducing the term *A-tuple* for an isolated a-tuple with $a \in A$, we may state the next two interpretations:

(i) $N(n,r,s)$ gives the number of binary vectors in n-space with exactly r components equal to "1", containing s A-tuples

(ii) $N(n,s) = \sum_{r=0}^{n} N(n,r,s)$ gives the total number of binary vectors in n-space with s A-tuples.

The double generating function of the numbers $N(n,r,s)$ is given by (20), (21) or alternatively by (22). The generating function of the numbers $N(n,s)$, $n = s, s+1, \dots$ has the much simpler form (23). Note that, for $A = \{k\}$ we obtain

$$f(t) = t^k \qquad g(t) = \frac{1}{1-t} - t^k$$

and substituting in (23) yields

$$\sum_{n=s}^{\infty} N(n,s)x^n = F(1,x;s) = \frac{x^{sk+s-1}(1-x)^{s+1}}{(1-2x+x^{k+1}-x^{k+2})^{s+1}}$$
$$\sum_{n=1}^{\infty} N(n,0)x^n = F(1,x;0) - 1 = \frac{x(2-x^{k-1}+x^{k+1})}{1-2x+x^{k+1}-x^{k+2}}$$

which agree with Theorem 13 of Apostol (1988). These expressions could also be used for the derivation of the recurrences (given also by Apostol (1988)) for the numbers $N(n,0)$. Finally, it may be noted that, employing formulas (17) and

(25), we may also derive the generating function and recurrence relations for the cyclic analogue of the abovementioned problem (which seems not to have been studied yet).

d. Fibonacci and Lucas numbers

Consider first $A = \{2, 3, \ldots\}$. Then the corresponding quantities $N(n, r, s)$, $N_c(n, r, s)$ may be interpreted as the number of r-combinations of the n consecutive integers $\{1, 2, \ldots, n\}$ placed on a line and a circle respectively, with s successions of at least two consecutive integers. The corresponding generating functions (which imply directly certain reccurrences) are given again by Theorems 4 and 5.

In the special case $s = 0$, the problem reduces to the much simpler situation where no two integers consecutive are allowed to enter in the combination. The corresponding numbers $N(n, r, 0)$ are related to the famous "Problème de Ménages" which was solved by Kaplansky (1943). We mention here that it is very easy to derive an explicit expression for $N(n, r, 0)$ (and by (26), for $N_c(n, r, 0)$) by making use of Theorem 4. Thus, replacing $g(t) = 1 + t$ in (22) we obtain

$$F(t, x; 0) = \frac{1 + tx}{1 - x(1 + tx)}$$

and expanding the right hand side into powers of t and x we deduce the well known formula

$$N(n, r, 0) = \binom{n - r + 1}{r}.$$

Notice also that the numbers

$$F_n = N(n - 1, 0) = \sum_{r=0}^{n-1} N(n - 1, r, s) \qquad n > 1$$

whose generating function is, by virtue of (23),

$$\sum_{n=1}^{\infty} F_n x^n = xF(1, x; 0) = \frac{x^2(x + 2)}{1 - x - x^2}$$

are the Fibonacci numbers.

Let us consider now $A = \{k, k + 1, \ldots\}$ with k a positive integer. Then the Theorem of Paragraph 3 in Hwang and Yao (1991), can be viewed as the special case $s = 0$ of (26), while Theorem 1 of Hwang (1986) can be extracted from the formula (see (22))

$$F(t, x; 0) = \frac{1 - t^k x^k}{1 - t - x(1 - t^k x^k)}$$

by expanding the right hand side in a double power series. For the corresponding cyclic case we may easily verify (see Corollary 4) that

$$\sum_{n=1}^{\infty} N_c(n,0)x^n = F_c(1,x;0) = x\frac{xg'(x)+g(x)}{1-xg(x)} = \frac{\sum_{j=1}^{k} jx^j}{1-\sum_{j=1}^{k} x^j}$$

which agrees with formula (2.10) of Charalambides (1991). The numbers $N_c(n,0)$ are known as Lucas numbers of order k, and have been extensively studied by Charalambides (1991).

e. Reliability of consecutive k-out-of-n systems

A consecutive k-out-of-n system is usually defined as a system of n components arranged on a line (linear consecutive k-out-of-n system) or a circle (circular consecutive k-out-of-n system) where the system fails if and only if at least k consecutive components fail. We consider here the case of independent components with common failure probability $q = 1 - p$. Introducing $A = \{k, k+1, \ldots\}$ it is not difficult to verify that the reliability of the linear system is

$$R(p,n) = \sum_{r=0}^{n} N(n,r,0)q^r p^{n-r}$$

while, for the circular system,

$$R_c(p,n) = \begin{cases} \sum_{r=0}^{n-1} N_c(n,r,0)q^r p^{n-r} & \text{if } n \geq k, \\ \sum_{r=0}^{n-1} N_c(n,r,0)q^r p^{n-r} + q^n & \text{if } 0 < n < k. \end{cases}$$

Hence

$$\sum_{n=0}^{\infty} R(p,n)z^n = \sum_{n=0}^{\infty}\left\{\sum_{r=0}^{n} N(n,r,0)q^r p^{n-r}\right\}z^n = F\left(\frac{q}{p}, pz; 0\right)$$

$$\begin{aligned} \sum_{n=1}^{\infty} R_c(p,n)z^n &= \sum_{n=1}^{\infty}\left\{\sum_{r=0}^{n-1} N_c(n,r,0)q^r p^{n-r}\right\}z^n + \sum_{n=1}^{\infty}(qz)^n \\ &= F_c\left(\frac{q}{p}, pz; 0\right) + qz\frac{1-(qz)^{k-1}}{1-qz} \end{aligned}$$

and making use of Theorems 4 and 5 we easily deduce the following expressions for the generating functions of the reliabilities of the linear and circular consecutive k-out-of-n systems

$$\sum_{n=0}^{\infty} R(p,n)z^n = \frac{1-(qz)^k}{1-z+pq^kz^{k+1}},$$

$$\begin{aligned}\sum_{n=0}^{\infty} R_c(p,n)z^n &= \frac{1-kpq^kz^{k+1}}{1-z+pq^kz^{k+1}} - \frac{(qz)^k}{1-qz} \\ &= \frac{1}{1-z+pq^kz^{k+1}}\left\{1-q^kz^k-(k-1)pq^kz^{k+1}+\sum_{n=k+2}^{2k} pq^{n-1}z^n\right\}.\end{aligned}$$

Expanding the right hand side of these formulas in power series with respect to z, one could readily obtain the expressions given by Lambiris and Papastavridis (1985) (Theorems 1 and 2), or Hwang (1986) (Theorem 4). Also, the recurrence relations mentioned by Lambiris and Papastavridis (1985) and Du and Hwang (1988) are immediately derived if we multiply the above formulas by $1-z+pq^kz^{k+1}$ and consider the coefficients of z^n in both sides. Note, also, the common denominator in the generating functions, which explains the fact that in both linear and circular systems the reliabilities satisfy the same recurrence relations for $n > 2k$.

References

Agin, M. and Godbole, A. (1991). A new exact runs test for randomness. *Computing Science and Statistics: Proc. 22nd Symposium on the Interface* (Edited by R. LePage and C. Page), 281–285, Springer-Verlag, New York.

Apostol, T. (1988). Binary vectors with prescribed subsets of consecutive ones. *J. Combin. Theory Ser.A* **47**, 176–190.

Bollinger, R. C. (1985). Strict consecutive-k-out-of-n: F systems. *IEEE Trans. Reliab.* **34**, 50–52.

Charalambides, Ch. (1991). Lucas numbers and polynomials of order k and the length of the longest circular success run. *Fibonacci Quarterly* **29**, 290–297.

Derman, C., Lieberman, G. J. and Ross, S. M. (1982). On the consecutive-k-of-n: F system. *IEEE Trans. Reliab.* **31**, 57–63.

Du, D. Z. and Hwang, F. K. (1988). A direct algorithm for computing reliability of a consecutive-k cycle. *IEEE Trans. Reliab.* **37**, 70–72.

Fu, J. C. (1993). Poisson convergence in reliability of a large linearly connected system as related to coin tossing. *Statistica Sinica* **3**, 161–175.

Godbole, A. P. (1990). Specific formulae for some success run distributions. *Statist. Probab. Lett.* **10**, 119–124.

Goulden, I. P. and Jackson, D. M. (1983). *Combinatorial Enumeration.* John Wiley, New York.

Goulden, I. P. and Jackson, D. M. (1984). A logarithmic connection for circular permutation enumeration. *Studies in Appl. Math.* **70**, 121–139.

Hirano, K. (1986). Some properties of the distributions of order k. *Fibonacci Numbers and their Applications: Proc. 1st Internat. Conf. on Fibonacci Numbers and their Applications* (Edited by Philippou, A. N., Bergum, G. E. and Horadam, A. F.), 43–53, Reidel, Dordrecht.

Hwang, F. K. (1981). Cycle polynomial. *Proc. Amer. Math. Soc.* **83**, 215–219.

Hwang, F. K. (1986). Simplified reliabilities for consecutive k-out-of-n systems. *SIAM J. Alg. Disc. Math.* **7**, 258–264.

Hwang, F. K. and Yao, Y. (1991). A direct argument for Kaplansky's theorem on cyclic arrangement and its generalization. *Oper. Res. Lett.* **10**, 241–243.

Kaplansky, I. (1943). Solution to the "Problème des Ménages". *Bull. Amer. Math. Soc.* **49**, 784–785.

Konvalina, J. (1981). On the number of combinations without unit separation. *J. Combin. Theory Ser.A* **31**, 101–107.

Lambiris, M. and Papastavridis, S. G. (1985). Exact reliability formulas for linear and circular consecutive-k-out-of-n: F systems. *IEEE Trans. Reliab.* **34**, 124–126.

Ling, K. D. (1989). A new class of negative binomial distributions of order k. *Statist. Probab. Lett.* **7**, 371–376.

Moser, W. and Abramson, M. (1979). Enumeration of combinations with restricted differences and cospan. *J. Combin. Theory* **7**, 162–170.

Papastavridis, S. G. (1990). m-consecutive-k-out-of-n: F systems. *IEEE Trans. Reliab.* **39**, 386–388.

Philippou, A. N. and Makri, F. S. (1986). Successes, runs and longest runs. *Statist. Probab. Lett.* **4**, 101–105.

Riordan, J. (1958). *An Introduction to Combinatorial Analysis.* John Wiley, New York.

Department of Mathematics, University of Athens, Panepistemiopolis, Athens 15784, Greece.

(Received April 1991; accepted December 1992)