# The Number Of Certain k-Combinations Of An n-Set<sup>\*</sup>

Thomas Wieder<sup>†</sup>

Received 1 March 2007

#### Abstract

We will count the number of possible coalitions. In combinatorial terms we will count the number of k-combinations formed from an n-set under certain restrictions. In contrast to the usual definition of a coalition in quantitative sociology, our k-combination needs not to cover the entire set. We will discern among disjoint and conjoint k-combinations and among those with or without the empty set and the n-set itself as allowed subsets. Several relations to and among certain integer sequences will be outlined.

### 1 Introduction

In the literature on multi-agents behaviour, a coalition is commonly defined as a subset U of a set S and a coalition structure as a partition of S into disjoint subsets. The empty set is not included into the common definition. For example, in a multi-agent system composed of three agents  $\{a_1, a_2, a_3\}$ , there exist seven possible coalitions:  $\{a_1\}, \{a_2\}, \{a_3\}, \{a_1, a_2\}, \{a_1, a_3\}, \{a_2, a_3\}, \{a_1, a_2, a_3\}$  and five possible coalition structures:  $\{\{a_1\}, \{a_2\}, \{a_3\}, \{a_1\}, \{a_2, a_3\}, \{\{a_2, a_3\}, \{\{a_2, a_3\}, \{\{a_2\}, \{a_3, a_1\}\}, \{\{a_3\}, \{a_1, a_2, a_3\}, \{a_1, a_2, a_3\}\}, \{\{a_2, a_3, a_1\}\}, \{\{a_3\}, \{a_1, a_2, a_3\}, \{a_1, a_2, a_3\}, \{a_1, a_2, a_3\}, \{a_1, a_2, a_3, a_1\}, \{a_1, a_2, a_3, a_1\}, \{a_1, a_2, a_3\}, \{a_1, a_2, a_3\}, \{a_1, a_2, a_3, a_1, a_2\}, \{a_1, a_2, a_3\}, \{a_1, a_2, a_3\}, \{a_1, a_2, a_3, a_1\}, \{a_1, a_2, a_3, a_1, a_2\}, \{a_1, a_2, a_3\}, [1]$ . What is meant here is a complete set partitioning, that is all n elements of S are included. Then the number of coalition structures for given n is equal to the Bell number B(n). According to our definition given below, a k-combination or equivalently a coalition structure needs not to cover all n elements. This approach may be more natural, if we think of a large society with correspondingly large n. Usually, not all of the n society members will take part in a coalition structure. Therefore, we arrive at another counting of coalition structures.

Our counting resides on the combinatorial side, but a complete practical point of view can be derived from the political side. For details, readers can refer to [2].

Consider the *n*-set  $S = \{1, 2, 3, ..., n\}$  composed of the labelled elements 1, 2, ..., e, ..., n. As well-known, there are  $2^n$  subsets  $U_s$ ,  $s = 1, ..., 2^n$  of S. The set

$$D(n,k) = \{U_{s=1}, U_{s=2}, \dots, U_{s=k}\}$$

formed by k disjoint subsets  $U_s \neq \emptyset \subset S$ , is called a disjoint strict k-combination. Two disjoint subsets do not have elements in common. DC(n, k) denotes the number of

<sup>\*</sup>Mathematics Subject Classifications: 05A18

<sup>&</sup>lt;sup>†</sup>Kassel, Germany, email: thomas.wieder@t-online.de

disjoint strict k-combinations. The set

$$D'(n,k) = \{U_{s=1}, U_{s=2}, \dots, U_{s=k}\}$$

formed by k disjoint subsets  $U_s \subseteq S$  (and  $\emptyset$  is allowed), is called a disjoint usual k-combination. DC'(n, k) denotes the number of disjoint usual k-combinations.

The set  $K(n,k) = \{U_{s=1}, U_{s=2}, \ldots, U_{s=k}\}$  formed by k subsets  $U_s \neq \emptyset \subset S$ , is called a conjoint strict k-combination. Two conjoint subsets may have elements in common. KC(n,k) denotes the number of conjoint strict k-combinations. The set  $K'(n,k) = \{U_{s=1}, U_{s=2}, \ldots, U_{s=k}\}$  formed by k subsets  $U_s \subseteq S$ , is called a conjoint usual k-combination. KC'(n,k) denotes the number of conjoint usual k-combinations.

From above definitions, we note that the difference between the strict case and the usual case is whether the empty set  $\emptyset$  and the set S are included. The k-combinations for k = 0 and k = n correspond to the empty set  $\emptyset$  and the entire set S.

### 2 Strict *k*-Combinations

#### **2.1** Disjoint strict k-Combinations $D(n,k)_t$

At first, we will calculate the number DC(n, k) of disjoint strict k-combinations. To construct a k-combination we need to select at least k elements from S, this can be done in C(n, k) ways where C(n, k) is the binomial coefficient. We even may take up to n elements which leaves us with a sum over the binomial coefficients C(n, i) running from i = k to n. But any of such a selection of i elements must be partitioned into k disjoint subsets. As well known, this can be done in  $S_2(i, k)$  ways. Thus we arrive at a sum over a product of binomial coefficients with Stirling numbers of the second kind and we have with  $n, k = 0, 1, 2, 3, \ldots$  and  $k \leq n$ 

$$DC(n,k) = \sum_{i=k}^{n} C(n,i)S_2(i,k).$$
 (1)

Eq. (1) is nothing but a new recurrence for the Stirling numbers of the second kind, namely  $S_2(n + 1, k + 1) = \sum_{i=k}^{n} C(n, i)S_2(i, k)$ . The following proof for the new recurrence has been given by referee 1 of this paper: We can count DC(n, k) in the following way. We divide S into two disjoint parts  $S_1$  and  $S_2$ .  $S_1$  needs to be partitioned into k disjoint subsets  $S_1^1, \ldots, S_1^k$ .  $S_2$  is a block, into we put 0 as label. Then  $S_1^1, \ldots, S_1^k, S_2 \cup \{0\}$  is a k + 1 partition of  $S \cup \{0\}$ . Therefore, we have

$$DC(n,k) = S_2(n+1,k+1).$$
(2)

For more information on the Stirling numbers of the second kind  $S_2(n, k)$  please refer to sequence A008277 of the OEIS [3]. With k = 2 we receive from (2) for n = 2, 3, 4, ...the sequence A000392 [3], that is 1, 6, 25, 90, 301, 966, 3025, 9330. These k-combinations are a good example for exploding combinatorial structures, since DC(n, k) grows rapidly. For small n, it is possible to construct the k-combinations explicitly by a computer program. It is instructive to see an example. For n = 3 and k = 2 we have  $\{\{1\}\{2\}\}, \{\{2\}\{3\}\}, \{\{1\}\{3\}\}, \{\{12\}\{3\}\}, \{\{1\}\{23\}\}, \{\{13\}\{2\}\}\}$ . Since we deal with sets only, the order of the elements within each subsets as well the order of the subset within a k-combination does not matter.

Just to get an impression, let us look at an assembly of n = 4 persons like in a family. This family may be split up into k = 2 coalitions debating their next destination for vacancies. Then we can have  $\sum_{i=2}^{4} C(4, i) S_2(i, 2) = 25$  different k = 2-coalition structures or, to put it more dramatically k = 2-confrontation structures. In many of these cases, several family members will not belong to any of the confronting parties (since the coalition structure needs not to contain all members). We can speak of neutral members and hopefully they will appease the conflict.

If we sum over k = 0, ..., n in (2), then we receive the total number of possible k-combinations for given n. This sum is well-known, it is just the Bell number B(n+1). Then, we can form an expression for the probability W(n, k) of a k-combination for given n

$$W(n,k) = \frac{S_2(n+1,k+1)}{B(n+1)}.$$
(3)

As an example, the probability for a k = 4-combination among n = 10 is  $W(n = 10, k = 4) = S_2(11, 5)/B(11) = 24673/67857 \approx 0.36$ . In more pictorial words, in a group formed by n = 10 "discernible members" = "individuals" we have a chance of 0.36 to meet with k = 4 coalitions. Not all of these individuals need to belong to one of these coalitions. Amazingly, there are  $S_2(n + 1 = 11, k + 1 = 5) = 24673$  different realizations of such a k = 4-party structure.

Let

$$D(n,k) = \{D(n,k)_{t=1}, \dots, D(n,k)_{t=DC(n,k)}\}$$

be the set of all possible k-combinations  $D(n,k)_t$  for given n and k. Eq. (1) allows us to calculate the total number  $N_e^D(n,k)$  of elements e in D(n,k). We just need to multiply the sum term by i and thus have

$$N_e^D(n,k) = \sum_{i=k}^n C(n,i) S_2(i,k)i.$$
 (4)

For k = 1 we meet with sequence A001787 which is equal to  $n \ 2^{(n-1)}$  and we actually have  $\sum_{i=k}^{n} C(n,i)S_2(i,1) \ i = \sum_{i=1}^{n} C(n,i) \ i = A001787(n)$ , see formula section in [3]. For k = 2 we have for  $n = 2, 3, 4, \ldots$  the numbers  $N_i = 2, 15, 76, \ldots$  and in fact we have 15 elements in the example above.

A further quantity can be derived by (4) and (2), that is the average number  $N_e^{D_t}(n,k)$  of elements  $1, 2, \ldots, e, \ldots, n$  within a k-combination  $D(n,k)_t$  for given n. It is

$$N_{e}^{D_{t}}(n,k) = \frac{N_{e}^{D}(n,k)}{DC(n,k)}.$$
(5)

#### **2.2** Conjoint strict k-combinations $K(n,k)_t$

If we relax the condition on the disjointness of the subsets  $U_s$ , then we arrive at the possibly (not necessarily) conjoint strict k-combinations K(n, k). Their numbers

||K(n,k)|| will be designated by KC(n,k). It is easy to calculate KC(n,k) since any combination of subsets is allowed, just the sets  $\emptyset$  and U = S are excluded. Then we have with  $n, k = 0, 1, 2, 3, \ldots$  and  $k \leq n$ 

$$KC(n,k) = C(2^n - 2, k).$$
 (6)

If n = 3 and k = 2, then we have

 $\{\{1,3\},\{1,2\}\},\{\{3\},\{2\}\},\{\{3\},\{1,3\}\},\{\{1\},\{2,3\}\},\{\{1\},\{1,2\}\},\{\{1\},\{1,3\}\},\{\{1\},\{3\}\},\{\{1,3\},\{2\}\},\{\{1,3\},\{2,3\}\},\{\{3\},\{2,3\}\},\{\{3\},\{1,2\}\},\{\{2,3\},\{2\}\},\{\{1,2\},\{2,3\}\},\{\{1,2\},\{2,3\}\},\{\{1\},\{2\}\}\}.$ 

The first values of KC(n, k = 2) are KC(2, 2) = 1, KC(3, 2) = 15, KC(4, 2) = 91, KC(5, 2) = 435, KC(6, 2) = 1891. These numbers have the form i(2i - 1) and the values for *i* follow from  $i = 2^{(n-1)} - 1$ . The numbers KC(n, 2) are thus hexagonal numbers (A000384) for  $i = 1, 3, 7, 15, 31, \ldots$  and another formula for them therefore is  $KC(n, 2) = 4^n/2 - 5 \cdot 2^{(n-1)} + 3$  [3].

Commonly, a person does not belong to two or more parties. Party membership is exclusionary, but club membership is not. So, a person may belong to the local football club and to the country's most famous football club, too. We could say that the conjoint k-combinations describe the club structure of S. In the strict case, both a party and a club which contains all n individuals is prohibited.

### 2.3 Difference KC(n,k) - DC(n,k)

The difference

$$KC(n,k) - DC(n,k) = C(2^n - 2, k) - S_2(n+1, k+1)$$
(7)

leads to the number of Boolean functions. For k = 2 we get the sequence A051375 [3], that is "the number of Boolean functions of n variables and rank 3 from Post class  $F(5, \inf)$ ", see the reference cited for definitions and literature on Boolean functions. Several equations are given in the reference and it is an open task to show the equivalence of (7) to them. Just for a comparison with a similar equation given in section 3.3 we cite here that  $A051375(n) = (1/2!)(4^n - 3^n - 3 \cdot 2^n + 5)$ .

### **3** Usual *k*-Combinations

We will use the ' to indicate a usual k-combination in contrast to a strict one. In a usual k-combination the sets  $\emptyset$  and S are allowed subsets.

### **3.1** Disjoint usual k-combinations $D'(n,k)_t$

To construct a usual disjoint k-combination  $D'(n, k)_t$  we combine k subsets  $U_s$  from  $S = \{1, 2, 3, ..., n\}$  which must be mutually disjoint, but we include the empty set  $\emptyset$  and S itself. The number of usual disjoint k-combinations will be written as DC'(n, k).

Again using a computer program we examined the first cases for small n. As an example, for n = 3 and k = 2 we have

$$\{ \emptyset, \{3\} \}, \{ \emptyset, \{1,3\} \}, \{ \emptyset, \{1,2,3\} \}, \{ \emptyset, \{2,3\} \}, \{ \emptyset, \{2\} \}, \{ \emptyset, \{1,2\} \}, \{ \emptyset, \{1\} \}, \\ \{ \{3\}, \{2\} \}, \{ \{1\}, \{2\} \}, \{ \{1\}, \{3\} \}, \{ \{1\}, \{2,3\} \}, \{ \{3\}, \{1,2\} \}, \{ \{1,3\}, \{2\} \}.$$

The observed sequence with n = 1, 2, 3, ... and k = 2 is 1, 4, 13, 40, 121, 364 and coincides with  $(3^n - 1)/2$  (A003462) [3]. In fact this is the Gaussian binomial coefficient  $G[n, 1]_{q=3}$ .

The number DC'(n, k) of disjoint usual k-combination includes the number DC(n, k) of disjoint strict k- combinations since the corresponding disjoint strict k-combinations are included into K(n, k). Further k- combinations arise from disjoint combinations of  $\emptyset$  with subsets from S. The problem is to count only disjoint k-combinations. For that purpose we can consider  $\emptyset$  as a further element 0 and thus go from  $S = \{1, 2, 3, ..., n\}$  to  $S' = \{0, 1, 2, 3, ..., n\}$ . The number of disjoint strict k-combinations for S' is from eq. (2) just  $S_2(n+1, k)$ . Then we have in total with n, k = 0, 1, 2, 3, ..., and  $k \le n$ 

$$DC'(n,k) = S_2(n+1,k+1) + S_2(n+1,k).$$
(8)

We observe that with k = 1 we have  $S_2(n+1, k+1) + S_2(n+1, k) = 2^n$ , that is the number of subsets of S. This circumstance makes perfectly sense with our definition of a k-combination. The subsets of S are the usual k-combination for S with k = 1. The same sum with k = n yields in a well-known formula for the central polygonal numbers (A000124) [3] which are 1, 2, 4, 7, 11, 16, 22, .... Thus the central polygonal numbers give the number of disjoint k-combinations of S with k = n.

#### **3.2** Conjoint usual k-combinations $K'(n,k)_t$

For a conjoint usual k-combination  $D'_t$  we combine k subsets  $U_s$  from  $S = \{1, 2, 3, ..., n\}$  which may be disjoint, but we include the empty set  $\emptyset$  and S itself. The number of usual conjoint k-combinations will be written as KC'(n,k). Our computer program provides us with an example for n = 3, k = 2 for which KC'(3, 2) = 28:

$$\begin{split} &\{\{1\},\{2\}\},\{\{2\},\{3\}\},\{\{1,2\},\{1,3\}\},\{\{3\},\{1,3\}\},\{\emptyset,\{2,3\}\},\{\{1,2,3\},\{2\}\},\\ &\{\emptyset,\{1,2,3\}\},\{\{1,2,3\},\{1\}\},\{\{1,2,3\},\{3\}\},\{\{1,2,3\},\{2,3\}\},\{\{3\},\{1,2\}\},\\ &\{\{2\},\{2,3\}\},\{\{1,2,3\},\{1,2\}\},\{\emptyset,\{3\}\},\{\emptyset,\{1,3\}\},\{\emptyset,\{2\}\},\{\{2\},\{1,3\}\},\\ &\{\{1\},\{2,3\}\},\{\{1\},\{1,3\}\},\{\{1\},\{1,2\}\},\{\{1,2,3\},\{1,3\}\},\{\{1\},\{3\}\},\{\emptyset,\{1\}\},\\ &\{\{1,3\},\{2,3\}\},\{\emptyset,\{1,2\}\},\{\{2\},\{1,2\}\},\{\{1,2\},\{2,3\}\},\{\{3\},\{2,3\}\}. \end{split}$$

In order to calculate KC'(n, k) we just have to form all  $2^n$  combinations which are possible for given n. Then we select k of them, we need not to cancel any of these  $C(2^n, k)$  combinations and arrive with  $n, k = 0, 1, 2, 3, \ldots$  and  $k \leq n$  at

$$KC'(n,k) = C(2^n,k).$$
 (9)

KC'(n,k) growths rapidly with n. In the case of k = 2 we find 1, 6, 28, 120, 496, 2016, 8128, 32640,... which is sequence A006516 [3]. From the formulas given in

A006516 we conclude without proof that  $KC'(n,2) = 2^{(n-1)}S_2(n+1,2)$  which gives the connection to eq. (8). Furthermore we conclude without proof from A006516 [3] that  $KC'(n,2) = G[n+1,2]_{q=2}G[n,2]_{q=2}$ . Again we meet with the Gaussian binomial coefficients as in the previous section 3.1.

### 3.3 Difference KC'(n,k) - DC'(n,k)

If we do the subtraction

$$KC'(n,k) - DC'(n,k) = C(2^n,k) - S_2(n+1,k+1) - S_2(n+1,k),$$
(10)

then we gain another class of integer sequences. The cases for k = 2 and k = 3 have been described already, see A036239 [3] and A036240 [3]. Interestingly, the interpretation given to them is "number of k-element intersecting families of an n- element set; number of k-way interactions when k subsets of power set on  $\{1...,n\}$  are chosen at random". By construction the eqs. (7) and (10) are closely related and for comparison with 3.3 we cite the explicit equation for k = 2 which is  $A036239(n) = (1/2!)(4^n - 3^n - 2^n + 1)$ .

### 4 Strict *k*-Combinations for Unlabeled Elements

Consider a set  $S^* = \{*, *, ..., *\}$  of n unlabeled elements  $*, |S^*| = n$ . Since we can not discern among the elements we do not ask for disjoint or conjoint subsets of  $S^*$ . We just select k subsets  $U_s$  of  $S^*$  and combine them to a set  $D^*(n,k)_t$  of k subsets under the condition that the number of elements of  $D^*(n,k)_t$  does not exceed the number of elements of  $S^*$ . We call the set  $D^*(n,k)_t$  a strict k-combination if we do not allow the empty set  $\emptyset$  and  $S^*$  itself as subsets. The total number  $||D(n,k)_t||^{e,*}$  of elements \* in  $D^*(n,k)_t$  needs not to equal n, but it is  $||D(n,k)_t||^{e,*} \leq n$ . As always, we give an example, namely D(n = 6 and k = 3):  $\{\{*\}, \{*\}, \{*\}, \{*\}, \{\{\{*\}, \{\{\{+\}, \{\{*\}, \{\{\{+\}, \{\{\{+\}, \{\{+\}, \{\{\{+\}, \{\{+\}, \{\{\{+\}, \{$ 

In the unlabelled case the number  $K(n, k)^*$  of strict k-combinations follows from the fact that the k-combinations base on partitions of integers. For given n and k we start at i = k, form all integer partitions of i into k parts, and then go on to i = k + 1and so on until i = n. However, the case k = 1 and simultaneously i = n selects  $S^*$  which is excluded for a strict k-combination. If we want to be strict too, then we exclude the case k = 1 and i = n from the following formula. If P(i, k) denotes the number of integer partitions of i into k parts then we have with  $n, k = 0, 1, 2, 3, \ldots$  and  $k \leq n$ 

$$K^{*}(n,k) = \sum_{i=k}^{n} P(i,k).$$
(11)

If one inspects  $K(n, k)^*$  with n = 1, 2, 3, ... for k as parameter then one gets for k = 2 the well-known sequence "quarter-squares"  $= floor(n/2) \cdot ceiling(n/2)$ , that is A002620 [3], for k = 3 one gets the expansion of  $1/((1-x)^2(1-x^2)(1-x^3)) = A000601$  [3], and for k = 4 one receives the expansion of  $1/((1-x)^2(1-x^2)(1-x^3)(1-x^4)) = A100262$  [3]. In particular A002620 has many combinatorial interpretations, here we

have added another. From these sequences we conclude without proof on the ordinary generating function o.g.f. for  $K(n,k)^*$ 

$$o.g.f = \frac{1}{(1-x)^2 \prod_{i=2}^{n} (1-x^i)}.$$
(12)

As in the labeled case, we can determine the number  $N_e^{D^*}(n,k)$  of all elements in  $D^*(n,k)$  for given n and k. This number is

$$N_e^{D^*}(n,k) = \sum_{i=k}^{n} P(i,k)i.$$
 (13)

In the example above with n = 6 and k = 3 we have 35 elements \* and indeed from (13) we have  $N_e^{D^*}(n,k) = 35$ .

In the unlabeled case the average number  $N_e^{D_t^*}(n,k)$  of elements from n elements \* within a k-combination of  $S^*$  is

$$N_e^{D_t^*}(n,k) = \frac{N_e^{D^*}(n,k)}{K(n,k)^*}.$$
(14)

For example, a k = 3-combination for n = 6 has on average  $N_e^{D_t^*}(n, k) = 5$  elements.

### 5 Usual *k*-Combinations for Unlabeled Elements

Now we form usual k-combinations from the set  $S^* = \{*, *, ..., *\}$  of n unlabeled elements \* which means we allow for the empty set  $\emptyset$  and the set  $S^*$  itself, too. The number  $K^{'*}(n,k)$  of these usual k-combinations includes the number K'(n,k)of strict k-combinations. Now k-combinations containing  $\emptyset$  accrue, their number is  $\sum_{i=k}^{n} P(i, k-1)$  since the k-th subsets is  $\emptyset$ . The case k = 1 and i = n selects  $S^*$  which is included now. In total we gain with  $n, k = 0, 1, 2, 3, \ldots$  and  $k \leq n$ 

$$K^{'*}(n,k) = \sum_{i=k-1}^{n} \left[ P(i,k) + P(i,k-1) \right].$$
(15)

With n = 0, 1, 2, 3, ... we get from (15) for k = 1 just the integers, but k = 2 is no more trivial and we get 0, 1, 3, 5, 8, 11, 15, ... which is OEIS sequence A024206 = "expansion of  $x^2(1 + x - x^2)/((1 - x^2)(1 - x)^2)$ " [3]. The OEIS provides us with the formula A024206(n) = A002620(n + 1) - 1 which is the connection between (15) and (11) in the case k = 1. With k = 3 we come to A034198 = "number of binary codes (not necessarily linear) of length n with 3 words" [3].

### 6 Conclusion

In looking for the possible coalitions among n persons, we allow for a partial set partition of the *n*-set S instead of a complete partition. This circumstance leads to the observation that the number of possible coalitions (disjoint strict k-combinations) among n persons is B(n + 1) instead of B(n) (see section 2.1, eq. (3)). While this observation does not look too exiting, it makes quite a difference theoretically and computationally. The computational difference can be best outlined by an example. For n = 4 we have B(4 + 1) = 52 possible coalitions instead of B(4) = 15 only. As a consequence, the probability W(n, k) to encounter a coalition of k persons within S is completely different numerically in comparison to a coalition definition based on a complete set partition. Furthermore, we have introduced the notion of conjoint k- combinations which could be used to describe inclusive social structures where a certain person may belong to several communities ("clubs") instead of one community ("coalition") only.

Acknowledgments. I am indebted to Neil J.A. Sloane for his patience and his generosity with my contributions to the On-Line Encyclopedia of Integer Sequences (OEIS). Many thanks to the referees for their constructive reports. Referee 1 provided a better proof for eq. (2). Last but not least many thanks to AMEN's Editor in Chief Prof. S. S. Cheng for his kindness!

## References

- V. D. Dang and N. R. Jennings, Generating coalition structures with finite bound from the optimal guarantees, International conference on Autonomous Agents and Multi-Agent Systems AAMAS'04, July 19-23, 2004, New York, USA.
- [2] G. Sartori, Parties and Party Systems: A framework for analysis, European consortium for political research, ECPR Press, 2005, ISBN 0954796616.
- [3] The Online Encyclopedia Of Integer Sequences, http://www.research.att.com/ ~njas/sequences/.