Independent subsets of a finite set

In probability theory the *probabilistic* or *stochastic* independence of two subsets A and B of some set Ω frequently is defined by the equation

$$P(A \cap B) = P(A) \cdot P(B) \tag{1.1}$$

where P is the probability function (probability measure) of the probability space under consideration.

In the following we will consider subsets A and B of a finite n-element set Ω . We will use $P(X) := \frac{|X|}{|\Omega|} = \frac{|X|}{n}$ as a "probability" function and define the subsets A and B to be *independent* if the following concretion of equation (**??**) holds:

$$\frac{|A \cap B|}{n} = \frac{|A|}{n} \cdot \frac{|B|}{n} \tag{1.2}$$

The sequence of the numbers satisfying this criterion of independence is discussed in the Online Encyclopedia of Integer Sequences OEIS A121312.

In OEIS A158345 the discussion is restricted to the case where A and B are **proper** subsets of Ω; i.e. $A \subsetneqq \Omega$ and $B \subsetneqq \Omega$.

A further interesting restriction might be the case where the subsets A and B are required to be **non-trivial**, i.e. $\emptyset \subsetneqq A \subsetneqq \Omega$ and $\emptyset \subsetneqq B \subsetneqq \Omega$ or, in other words, $1 \leq |A| \leq n-1$ and $1 \leq |B| \leq n-1$.

For the following let $a := |A|$, $b := |B|$ and $d := |A \cap B|$. Then the crucial equality for the independence of the sets A and B is

$$\frac{a}{n} \cdot \frac{b}{n} = \frac{d}{n} \tag{1.3}$$

or

$$a \cdot b = d \cdot n \tag{1.4}$$

Some consequences of equation (**??**):
* For $n = 0$ (i.e. $\Omega = \emptyset$) equation (**??**) is true.
* If n is prime, then in the "non-trivial subsets case" there are no independent subsets A and B.
* For a better chance for the independence of the subsets it is favorable if n has many divisors. If furthermore n is relatively small then the chance for the independence of subsets of Ω is even better.

Remark: In the following we consider the "non-trivial subsets case" (i.e. $1 \leq |A| \leq n-1$ and $1 \leq |B| \leq n-1$).
As is often the case, the problem can be tackled by a stochastic simulation for heuristic reasons. In the following figure, resulting from a simulation, the horizontal axis represents the cardinality n of Ω and the vertical axis represents the probability for randomly selected pairs of *non-trivial* subsets being independent. For details of the simulation see:
https://jochen-ziegenbalg.github.io/materialien/Manuskripte/Disko.pdf.

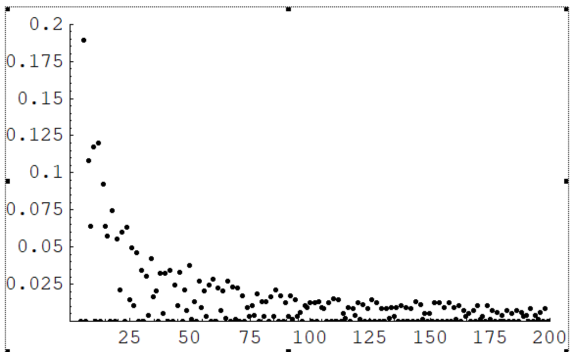

Figure 1.1: Siimulation with random subsets

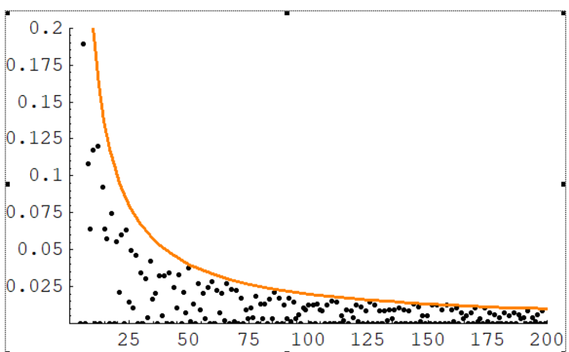

Figure 1.2: Simulation enhanced by the graph of $f(x) = \frac{2}{x}$

* Of course, the figures give only an approximative empirical picture. But certain observations are quite suggestive: The larger Ω is, the smaller, by trend, is the chance for the independence of two subsets of Ω.
* If n is prime then the simulation dots are on the x-axis. (See table **??**)
* The simulation dots are all below the graph of the function $f(x) = \frac{2}{x}$.

An *Example*: $n = 6, \quad \Omega = \{1, 2, 3, 4, 5, 6\}$. The power set of Ω is

$$
\begin{aligned}
\mathfrak{P}(\Omega) = \{ & \\
& \{\} \\
& \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \\
& \{1,2\}, \{1,3\}, \{1,4\}, \{1,5\}, \{1,6\}, \{2,3\}, \{2,4\}, \{2,5\}, \{2,6\}, \\
& \quad \{3,4\}, \{3,5\}, \{3,6\}, \{4,5\}, \{4,6\}, \{5,6\}, \\
& \{1,2,3\}, \{1,2,4\}, \{1,2,5\}, \{1,2,6\}, \{1,3,4\}, \{1,3,5\}, \{1,3,6\}, \\
& \quad \{1,4,5\}, \{1,4,6\}, \{1,5,6\}, \{2,3,4\}, \{2,3,5\}, \{2,3,6\}, \\
& \quad \{2,4,5\}, \{2,4,6\}, \{2,5,6\}, \{3,4,5\}, \{3,4,6\}, \{3,5,6\}, \{4,5,6\}, \\
& \{1,2,3,4\}, \{1,2,3,5\}, \{1,2,3,6\}, \{1,2,4,5\}, \{1,2,4,6\}, \{1,2,5,6\}, \\
& \quad \{1,3,4,5\}, \{1,3,4,6\}, \{1,3,5,6\}, \{1,4,5,6\}, \\
& \quad \{2,3,4,5\}, \{2,3,4,6\}, \{2,3,5,6\}, \{2,4,5,6\}, \{3,4,5,6\}, \\
& \{1,2,3,4,5\}, \{1,2,3,4,6\}, \{1,2,3,5,6\}, \{1,2,4,5,6\}, \{1,3,4,5,6\}, \\
& \quad \{2,3,4,5,6\}, \\
& \{1,2,3,4,5,6\} \ \ \}
\end{aligned}
\tag{1.5}
$$

There are 2^6 ($= 64$) subsets of Ω. Their cardinalities are given in Table **??**.

number of elements of the subset	number of subsets	formulation by formula
0	1	$\binom{6}{0} = \frac{6!}{0! \cdot 6!}$
1	6	$\binom{6}{1} = \frac{6!}{1! \cdot 5!} = \frac{6}{1}$
2	15	$\binom{6}{2} = \frac{6!}{2! \cdot 4!} = \frac{6 \cdot 5}{1 \cdot 2}$
3	20	$\binom{6}{3} = \frac{6!}{3! \cdot 3!} = \frac{6 \cdot 5 \cdot 4}{1 \cdot 2 \cdot 3}$
4	15	$\binom{6}{4} = \frac{6!}{4! \cdot 2!} = \frac{6 \cdot 5}{1 \cdot 2}$
5	6	$\binom{6}{5} = \frac{6!}{5! \cdot 1!} = \frac{6}{1}$
6	1	$\binom{6}{6} = \frac{6!}{6! \cdot 0!} = \frac{6}{6}$

Table 1: subsets of a 6-element set

For $n = 6$, dismissing the trivial subsets $\emptyset$ und Ω, there are 720 pairs of independent subsets - resulting in a quotient of $\frac{720}{(2^6-2) \cdot (2^6-2)} = \frac{720}{3844} \approx 0.1873$ (for a complete listing of the independent subsets see:

https://jochen-ziegenbalg.github.io/materialien/Manuskripte/Disko.pdf).

The following Maxima program gives the number of the independent subsets of an n-element set (compare with OEIS A121312).

```
independent_subsets(n) :=                                          (1.6)
  block([a, b, d, s : 2^(n+1)-1 ],
     /*  if  a=0  or  b=0  then independent;  2 * 2^n  cases */
     /*  1 subtracted for double counting case a=0 and b=0   */
    for d:1 thru n do
      for a:d thru n do
        (b : n*d / a,
         if integerp(b)
            then (s : s+binomial(n,a)*binomial(a,d)*binomial(n-a,b-d))),
    s ) ;
```

Example:
`independent_sets(6);`
`972`

The following list `L25` is an initial interval (from 0 to 25) of the integers. By use of the `map`-command of Maxima the corresponding numbers of independent sets are computed.

```
L25 : makelist(i, i, 0, 25);

map(independent_sets, L25);
[1, 4, 12, 28, 84, 124, 972, 508, 8020, 17164, 130092, 8188, 1794156,
32764,  23609052, 55986868, 274827860, 524284, 5338824444, 2097148,
63030243724, 117928401724, 995282568732, 33554428, 15265553226604,
14283226194724]
```

Remark: In program (**??**) the trivial subsets $\emptyset$ and Ω are included. By straightforward computation it can be shown that if A or B are equal to $\emptyset$ or to Ω, then criterion (**??**) for independence $a \cdot b = d \cdot n$ is satisfied.

If only the non-trivial subsets are to be considered then the following cases have to be omitted:

$(\emptyset, X)$ with $X \subseteq \Omega$; these are 2^n cases
$(Y, \emptyset)$ with $Y \subseteq \Omega$; these are 2^n Fälle
(Ω, U) with $U \subseteq \Omega$; these are 2^n Fälle
(V, Ω) with $V \subseteq \Omega$; these are 2^n Fälle

With this way of counting, however, the four pairs $(\emptyset, \emptyset)$, $(\emptyset, \Omega)$, $(\Omega, \emptyset)$, (Ω, Ω) are counted twice. Correcting this, the number of non-trivial independent subsets of an n-element set ist given by the following program:

$$Maxima - Program \tag{1.7}$$

```
independent_nontrivial_sets(n) :=
  if  is(n=0)  then 0 else independent_sets(n) - 4*(2^n -1);
```

Example:

```
independent_nontrivial_sets(6);
720
```

For the case of *proper* subsets the sequence OEIS A158345 should be considered. It is defined by: "The number of pairs of independent outcomes when rolling an n-sided die. Or in other terms, the number of pairs of proper subsets A, B of a set S, such that $\frac{|A|}{|S|} \cdot \frac{|B|}{|S|} = \frac{|A \cap B|}{|S|}$". Proper subsets of a set S are all subsets except S itself. Thus, if only the proper subsets are to be considered then the following cases have to be omitted:

(Ω, U) with $U \subseteq \Omega$; these are 2^n Fälle

(V, Ω) with $V \subseteq \Omega$; these are 2^n Fälle

With this way of counting, however, the pair (Ω, Ω) is counted twice. Correcting this, the number of proper independent subsets of an n-element set ist given by the following program:

```
independent_proper_sets(n) :=
  if  is(n=0)  then 0 else  independent_sets(n) - 2*2^n + 1  ;
```

In Table **??** some initial values of the discussed functions are shown:

n	independent pairs of all subsets	independent pairs of proper subsets	independent pairs of non-trivial subsets
0	1	0	0
1	4	1	0
2	12	5	0
3	28	13	0
4	84	53	24
5	124	61	0
6	972	845	720
7	508	253	0
8	8020	7509	7000
9	17164	16141	15120
10	130092	128045	126000
11	8188	4093	0
12	1794156	1785965	1777776
13	32764	16381	0
14	23609052	23576285	23543520
15	55986868	55921333	55855800
16	274827860	274696789	274565720
17	524284	262141	0
18	5338824444	5338300157	5337775872
19	2097148	1048573	0
20	63030243724	63028146573	63026049424
21	117928401724	117924207421	117920013120
22	995282568732	995274180125	995265791520
23	33554428	16777213	0
24	15265553226604	15265519672173	15265486117744
25	14283226194724	14283159085861	14283091977000
26	216345187553052	216345053335325	216344919117600
27	240143438812708	240143170377253	240142901941800
28	2854495035174300	2854494498303389	2854493961432480
29	2147483644	1073741821	0
30	55689700679133012	55689698531649365	55689696384165720]

Table 2: independent subsets of a finite n-element set