# Opgave B NAW 5/5 nr. 1 maart 2004

Jaap Spies

March 2004

## The problem

### Introduction

Consider the first digit in the decimal expansion of $2^n$ for $n \geq 0$: 1, 2, 4, 8, 1, 3, 6, 1, 2, 5, 1, 2, 4, $\cdots$. Does the digit 7 appear in this sequence? Which digit appears more often, 7 or 8? How many times more often?

### Solution

The first question is easily solved affirmative: $2^{46} = 70368744177664$.
The sequence is well known from Sloane's On-Line Encyclopedia of Integer Sequences as $A008952$. See [1].
We use the formula

$$a(n) = \lfloor 2^n / 10^{\lfloor n \cdot \frac{\ln 2}{\ln 10} \rfloor} \rfloor$$

in a Maple Program [2] to calculate the frequency of digit $d$ in all $a(k)$ with $k \leq n$.

```
 > A:=proc(n,d) local i,k,s;
   s:=0;
   for k from 0 to n do
    i:=floor(2^k / 10^floor(k*ln(2)/ln(10)));
    if i=d then s:=s+1 fi;
   od;
   RETURN(s);
  end;
> seq(A(500,i),i=1..9);
   151, 88, 63, 49, 39, 34, 28, 26, 23
```

Frequency count $f(d)$ for leading digit $d$ of $2^n$:

| $\leq n$ d | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 85 | 26 | 16 | 10 | 9 | 7 | 5 | 4 | 5 | 4 |
| 86 | 26 | 16 | 10 | 9 | 7 | 5 | 5 | 5 | 4 |
| 100 | 31 | 17 | 13 | 10 | 7 | 7 | 6 | 5 | 5 |
| 200 | 61 | 36 | 24 | 20 | 16 | 13 | 11 | 11 | 9 |
| 1000 | 302 | 176 | 125 | 97 | 79 | 69 | 56 | 52 | 45 |
| 2000 | 603 | 354 | 248 | 194 | 160 | 134 | 114 | 105 | 89 |
| 3000 | 904 | 529 | 374 | 291 | 238 | 201 | 173 | 155 | 136 |
| 4000 | 1205 | 705 | 499 | 388 | 317 | 269 | 230 | 207 | 181 |
| 5000 | 1506 | 882 | 623 | 485 | 397 | 335 | 288 | 259 | 226 |
| 6000 | 1807 | 1058 | 748 | 582 | 476 | 401 | 347 | 309 | 273 |
| 7000 | 2108 | 1233 | 874 | 679 | 554 | 468 | 406 | 359 | 320 |
| 8000 | 2409 | 1409 | 999 | 776 | 633 | 537 | 462 | 412 | 364 |
| 9000 | 2710 | 1587 | 1122 | 873 | 714 | 602 | 520 | 463 | 410 |
| 10000 | 3011 | 1761 | 1249 | 970 | 791 | 670 | 579 | 512 | 458 |
| Benford's law : | 3010 | 1761 | 1249 | 969 | 791 | 669 | 579 | 511 | 457 |

The second question can be answered by: '*the winner is* 7'! From $n > 209$ or some more, the frequency of the digit 7 is greater than that of 8. We can generalize: for $n$ large enough the frequencies are a decreasing sequence, meaning for digits $d_1$ and $d_2$: $d_1 < d_2$ implies $f(d_1) > f(d_2)$. We can think of a reason: multiplication of $2^n$ with leading digit 1 with $2^{10} = 1024$, gives more often the same leading digit, compared with the larger leading digits 2, 3, $\cdots$, 9, and so on.
But Benford's law comes to the rescue (see [3]), our sequence is a well known example:

$$\text{Prob(first significant digit} = d) \ = log_{10}(1 + \frac{1}{d}), \text{ for } d = 1, 2, ..., 9$$

The similarity of the last two lines in the table above is striking!
The last question is the most difficult to answer. Our best guess for large $n$ is according to Benford's law:

$$\frac{f(7)}{f(8)} \quad \text{tends to} \quad \frac{\log_{10}(1 + 1/7)}{\log_{10}(1 + 1/8)} = 1.133706496$$

# References

[1] http://www.research.att.com/~ njas/sequences/
[2] http://www.maplesoft.com/
[3] Hill, T.P., The Significant-Digit Phenomenon, Amer. Math. Monthly 102, 322-327, 1995