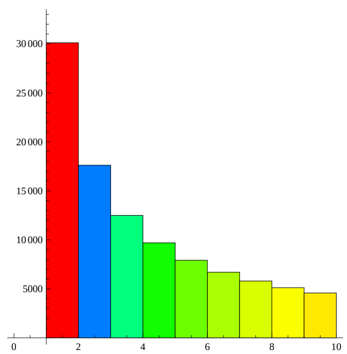
Lecture 13: Experimental mathematics

Experimental mathematics is doing mathematics close to physics in methodology. We experiment, for example with the help of a computer. Instead of talking about this abstractly, we will look during lecture at three examples illustrating this. The first is **Benford's law** which deals with the statistics of the first significant digit in data. Simon Newcomb found the law in 1881 and Frank Benford made significant progress on it in 1938. Here is an example where one can prove things. Look at the first digits of the sequence 2^n. One can prove that the digit k appears with probability $p_k = \log_{10}(1+1/k)$. The digit 1 for example occurs with about $\log_{10}(2) = 0.30$ which is 30 percent. Lets experiment and look at 2^n for $n = 1$ to $n = 100'000$ and determine the first digit:

```
data = Table[First[IntegerDigits[2^n]], {n, 1, 100000}];
S = Histogram[data, 10, ColorFunction -> Hue]
```

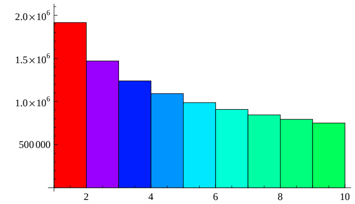


How does one compute the probability? If we look at the logarithms, then $\log(2^n) = n\log(2)$. The first digit is 1 if the rest of $[n\log(2)]$ modulo 1 is between 0 and $\log(2)$. The first digit is 2 if it is between $\log(2)$ and $\log(3)$ etc. The probability that the letter is k is $\log_{10}(k+1) - \log_{10}(k)$.

One can look at the first significant digit problem on other sequences like squares $1, 4, 9, 1, 2, 3, 4, 6, 8, 1, 1, 1, 1$
Here is an experiment:

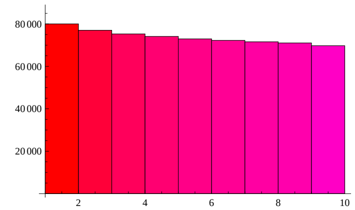
```
data = Table[First[IntegerDigits[n^2]], {n, 1, 10000000}];
S = Histogram[data, 10, ColorFunction -> Hue]
```

It is interesting because we want to see what the distribution of $2\log(n)$ is modulo 1. It looks as if we have a similar Benford law here. Indeed it is a generalized Benford law with $p_k = \frac{\int_k^{k+1} x^{-\alpha}\, dx}{\int_1^{10} x^{-\alpha}\, dx} = [(k+1)^{1-\alpha} - k^{1-\alpha}]/(10^{1-\alpha} - 1)$. It interpolates the Benford law $\alpha = 1$ with the uniform distribution $\alpha = 0$.
We have the digit 1, if $\log(n) \in k + [0, \log(2)]$. How many cases are in 1000 and 2000. It is $\sqrt{2000} - \sqrt{1000} = \sqrt{1000}(\sqrt{2} - 1)$. How many cases are in 2000 and 3000. It is $\sqrt{3000} - \sqrt{2000} = \sqrt{1000}(\sqrt{3} - \sqrt{2})$.



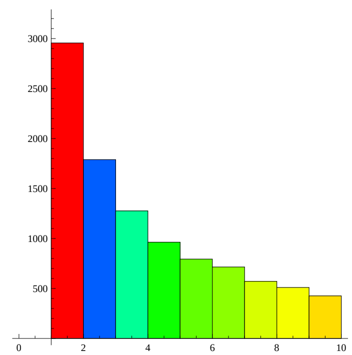
What is the first significant digit of the prime numbers?

```
data = Table[First[IntegerDigits[Prime[n]]], {n, 1, 664000}];
S = Histogram[data, 10, ColorFunction -> Hue]
```



How many primes are there in 1000 and 2000. We expect $1000/Log[1000]$ primes in there and $1000/Log[2000]$ with first significant digit 2.
S1=ListPlot[Table[PrimePi[k],k,10000]]; S2=ListPlot[Table[k/Log[k],k,10000]]; Show[S1,S2]
We expect the distribution to be $a/\log(k)$, where $a = \sum 1/\log(k)$.
For factorials, the limiting distribution is known to be the Benford distribution. There is no reason why $\log_{10}(n!) mod 1$ should not be uniformly distributed.

```
data = Table[First[IntegerDigits[n!]], {n, 1, 10000}];
S = Histogram[data, 10, ColorFunction -> Hue]
```



Also for the partition numbers, $p(n)$, which give the number of possibilities in which the number n can be written as a sum of integers, we measure that the Benford distribution takes place. As far as we know this is not known.

```
data = Table[First[IntegerDigits[PartitionsP[n]]], {n, 1, 10000}];
S = Histogram[data, 10, ColorFunction -> Hue]
```

s