# TNM098 – Lab 1

## Background

A list of 60 data points with coin tosses was given, named in the range 0-59. Each data point contained 200 tosses. The task was to find out which of the data points that were a random number generator and which ones that were humans faking randomness. In this report head and tails are represented with ones and zeros.

## Hypothesis

While a computer would create a string with a truly random result where the next toss is not dependent on any of the tosses before, a human would be biased by seeing the row of previous faked tosses. The human perception of randomness is different from what it really is. A real random string of tosses would contain sequences of consecutive heads or consecutive tails with a lot of different lengths. A human would feel that after about three heads in a row a tail is needed for it to look random enough. Also, since the human quickly shifts between heads and tails, the total number of heads or tails in a faked data point would be relatively close to 50 %.

## Method

First, all data points with a number of heads that was outside of the standard deviation were removed. They were most likely to be computers. Since coin toss is considered to be independent, it has a binomial distribution and the standard deviation is .

For the remaining data points, the length of sequences with consecutive heads or tails and the number of sequence of a certain length was counted. The values for sequences of 2, 3 and 5 consecutive digits were plotted in 2 dimensions, resulting in three plots that can be seen in table 1.

Table 1- Plots and the datapoints marked in each plot.

|  |  |
| --- | --- |
| **Plots** | **Data points** |
|  | 0, 1, 2, 3, 4, 5, 6, 7, 18, 20, 22, 26, 27, 39, 40, 46, 49, 50, 54, 59 |
|  | 0, 1, 5, 6, 7, 20, 22, 26, 27, 28, 32, 36, 37, 38, 39, 46, 48, 49, 50, 54 |
|  | 0, 1, 5, 7, 13, 15, 20, 22, 25, 26, 27, 36, 37, 38, 39, 48, 49, 50, 54, 58 |
| Mutual points in all images | 0, 1, 5, 7, 20, 22, 26, 27, 39, 49, 50, 54 |
| Mutual points in two images. Bold numbers were chosen to the result and marked with blue rings in images without them. | **6**, 36, **37**, **38**, 46, 48, |

The 20 points most likely to be humans in each plot were marked, and the mutual data points that were marked in every plot were decided to be human. The points that were marked in two plots were reviewed and three of them were a good fit in the points already marked and were therefore also chosen to be human.

## Result

After the elimination of point outside of the standard deviation 41 points were left. Points removed: 8, 9, 11, 16, 17, 19, 21, 23, 24, 29, 31, 33, 35, 42, 43, 45, 53, 56, 57.

The second part of the analysis resulted in the following data points to be humans (from 0 to 59):

0, 1, 5, 6, 7, 20, 22, 26, 27, 37, 38, 39, 49, 50, 54.