Probability Theory

When looking into probability we look for the likelihood of something occurring. It is measured on a scale of 0% to 100%, 0 to 1 or 0/1 to 1/1. In order to calculate the probability of an event you first need to find the total number of possible outcomes which is known as the set. When rolling a fair die the set would be the numbers appearing on the die [1, 2, 3, 4, 5 and 6]. As there are 6 entries in the set all calculations would need to be divided by 6. For example when calculating probability of rolling a six is 1/6 or 0.16. The probability of rolling two sixes in a row is 1/6 \* 1/6 which is equal to 1/36 or 0.027. Note we multiply the bottom part of the fraction, this is known as the ‘rule of the product’.

In order to determine the minimum length of sentence required for my PassSentence program I need to apply the above principals. When applying the same principals to selecting 12 possible letters for my PassSentence program, I first needed to determine the number of possible combinations of different letters that can be selected. This is done to avoid the same characters being used in a different order for a different password. In order to calculate the different number of combinations of the 12 letters I needed to compute the following calculation:

nCr = n! / r! (n - r)!

The ‘n!’ and ‘r!’ in the equation is notation for the factorial of n and r. “In mathematics, the factorial a non-negative integer is the product of all positive integers less than or equal to n multiplied together e.g. 5! = 5\*4\*3\*2\*1 = 120” (En.wikipedia.org, 2017). I chose to use a set (n) of 20 letters to begin with as I will be selecting 12 letters (r) from this set and 220 was a nice round number. The calculation went as follows:

**20C12 = 20! /12!(20-12)!**

**20! = 2,432,902,008,176,640,000**

**12! = 479,001,600**

**20-12 = 8**

**8! = 40320**

**2,432,902,008,176,640,000 / 479,001,600 \* 40320**

**2,432,902,008,176,640,000 / 19,313,344,512,000**

**= 125970**

This meant that the probability of selecting any combination of the 20 characters from the password is 1/125,970, which is equal to (approx.) 0.0000007. The probability of selecting 2 exact passwords when using a set of 20 letters is equal to 1/125,970 \* 1/125,970 = 1/15,868,440,900 =

(approx.) 0.000000000006

Therefore I have decided to choose a minimum length of 20 at this stage as the probability of the program selecting the same 12 letters on more than one occasion is so low. However, I may choose to increase this when I compute more probability aimed at the quantity of crossover of the selected numbers (how many times each letter appears in the total number of passwords).

1/5 1/10

References

En.wikipedia.org. (2017). *Factorial*. [online] Available at: https://en.wikipedia.org/wiki/Factorial [Accessed 26 Jan. 2017].