Assignment 2

Rubeena Aafreen (EE20RESCH11012)

Download all python codes from

https://github.com/rubeenaafreen20/EE5600AI−ML/tree/master/Assignment2/Code

and latex codes from

https://github.com/rubeenaafreen20/EE5600AI−ML/tree/master/Assignment2

1 Problem

On one page of a telephone directory, there were 200 telephone numbers. The frequency distribution of their unit place digit (for example, in the number 25828573, the unit place digit is 3) is given in Table below: Without looking at the page, the pencil is placed on one of these numbers, i.e., the number is chosen at random. What is the probability that the digit in its unit place is 6?

Digit	Frequency	Probability
0	22	0.11
1	26	0.13
2	22	0.11
3	22	0.11
4	20	0.1
5	10	0.05
6	14	0.07
7	28	0.14
8	16	0.08
9	20	0.1

TABLE 1: Given frequency distribution

2 Explanation

probability is defined as

$$P = \frac{\text{number of outcomes}}{\text{Sample space}} \tag{2.0.1}$$

3 Solution

Let $X \in \{i\}_{i=1}^{i=6}$ and f_i be the corresponding frequency. Then,

$$P_r(X = i) = \frac{f_i}{200} \tag{3.0.1}$$

From table 1,

$$P_r(X = 6) = \frac{14}{200} \tag{3.0.2}$$

$$= 0.07 \tag{3.0.3}$$

4 Output

The outputs of Python program are attached below:

Digit	Frequency	Probability
0	21	0.105
1	13	0.065
2	20	0.1
3	21	0.105
4	20	0.1
5	25	0.125
6	15	0.075
7	24	0.12
8	20	0.1
9	21	0.105

TABLE 2: For 200 randomly generated numbers

Digit	Frequency	Probability
0	1007	0.1007
1	988	0.0988
2	997	0.0997
3	1010	0.101
4	1005	0.1005
5	1018	0.1018
6	1000	0.1
7	984	0.0984
8	1019	0.1019
9	972	0.0972

TABLE 3: For 10000 randomly generated numbers

5 Explanation

The **Law Of Large Numbers** is a fundamental concept for probability and statistics. It states that that as the number of trials increase, the experimental probability will get closer and closer to the theoretical probability.
From the output tables 2 and 3, we can deduce that as the number of trials increase, the ratio of the number of successful occurrences to the number of trials will tend to approach the theoretical probability of the outcome for an individual trial. Since all the digits are equiprobable, ideally each probability should be 1/10=0.1
In table 3, when number of trials are 10,000, probability of each digit is approximately 0.1 with very little deviation. eg. 0.1005.
With 200 samples, Tables 2 and 3 are slightly different because the number of simulations is not sufficient for convergence in the probabilities.