Question 2 (7 marks)

The leading digit of a number is defined as the left-most non-zero digit. In the numbers given below, for example, 7 is the leading digit.

$$712 \qquad 74.2 \qquad 0.00798$$

Let X represent the leading digit of a naturally occurring number. Benford's law states that the leading digits in large sets of naturally occurring numbers are distributed as follows:

x	1	2	3	4	5	6	7	8	9
$\Pr(X = x)$	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

(a) What property of the values in the table above mean that X is **not** uniformly distributed?

(1 mark)

(b) Determine the probability of randomly selecting a naturally occurring number with a leading digit of 6 or greater.

(1 mark)

(c) Identify which of the following provides a correct calculation for the mean of the leading digit of a naturally occurring number, μ_X. Tick the appropriate box.

☐ $\mu_X = 1^2(0.301) + 2^2(0.176) + 3^2(0.125) + 4^2(0.097) + 5^2(0.079) + 6^2(0.067) + 7^2(0.058) + 8^2(0.051) + 9^2(0.046)$

☐ $\mu_X = 1(0.301) + 2(0.176) + 3(0.125) + 4(0.097) + 5(0.079) + 6(0.067) + 7(0.058) + 8(0.051) + 9(0.046)$

☐ $\mu_X = \dfrac{0.301 + 0.176 + 0.125 + 0.097 + 0.079 + 0.067 + 0.058 + 0.051 + 0.046}{9}$

(1 mark)

(d) Benford's law can be used to detect fraudulent financial records. Further investigation of a company's financial records is undertaken if a calculated 95% confidence interval supports the claim that the mean of the leading digit of the financial records is greater than 3.441.

Consider a random sample of 250 numbers from a company's records. The mean of the leading digit in this sample is $\bar{x} = 3.816$.

(i) Calculate a 95% confidence interval for the mean of the leading digit of the company's financial records, using the random sample described above. Assume that the standard deviation is $\sigma_X = 2.462$.

(2 marks)

(ii) Does this 95% confidence interval suggest that further investigation should be undertaken? Justify your answer.

(2 marks)