CS240, Spring 2022
Assignment 4: Question 4

Q4a) Which hash function is better? Justify your answer.

To compare the two hash functions, lets consider their values for the first 7 elements of our universe:

k	Output of h1(k)	Output of h2(k)
1	1	2
2	2	4
3	3	0
4	4	2
5	5	4
6	0	0
7	1	2

From this table we can see that the output of h1(k) encompasses the set {0,1,2,3,4,5} where as the output of h2(k) only encompasses {2,4,0}.

Therefore since the output of h2(k) only has half as many outputs as h1(k) we would expect more conflicts and thus h1(k) is the better hash function.

Q4b) Prove that $\lim_{n\to\infty} \frac{c_n}{n} = \frac{1}{e}$.

To start we know that the number of keys and the number of spaces in our hash table are both equal to n. Moving forward we will define an indicator variable I such that (where i is an integer between 0 and n-1):

$$I_i = \begin{cases} 0 & \text{if the hash table at index i has more then 0 elements chained in it} \\ 1 & \text{if the has table at index i has 0 elements in it} \end{cases}$$

Since c_n represents the expected number of empty elements in the hash table, therefore we get:

$$c_n = \sum_{i=0}^{n} I_i$$

An index will be empty if each of the n numbers that are inserted into the table go into any of the $n-1$ slots in the hash table. Therefore the expected value of any slot being empty is given by

$$E[I_i] = (\frac{n-1}{n})^n$$

Plugging this into our equation for c_n we get:

$$c_n = \sum_{i=0}^{n} (\frac{n-1}{n})^n$$
$$c_n = n(\frac{n-1}{n})^n$$

Since we are asked to solve for $\lim_{n\to\infty} \frac{c_n}{n}$ we can now plug our equation for c_n in to get:

$$\begin{aligned}
\lim_{n\to\infty} \frac{c_n}{n} &= \lim_{n\to\infty} \frac{n(\frac{n-1}{n})^n}{n} \\
&= \lim_{n\to\infty} (\frac{n-1}{n})^n \\
&= \lim_{n\to\infty} (1 - \frac{1}{n})^n \\
&= \frac{1}{e} \quad \text{from the fact given in the question}
\end{aligned}$$

Thus we have proved $\lim_{n\to\infty} \frac{c_n}{n} = \frac{1}{e}$ as required.