Problem 1 (CLRS 11.2-1). (1 point) Suppose we use a hash function \( h \) to hash \( n \) distinct keys into an array \( T \) of length \( m \). Assuming simple uniform hashing, what is the expected number of collisions? More precisely, what is the expected cardinality of \( \{ \{k, l\} : k \neq l \text{ and } h(k) = h(l)\} \)?

Problem 2 (CLRS 11.3-3). (3 points) One of the simplest hash functions can be constructed using the division method: \( h(k) = k \mod m \). Consider a version, where \( m = 2^p - 1 \) and \( k \) is a character string interpreted in radix \( 2^p \). Show that if we can derive string \( x \) from string \( y \) by permuting its characters, then \( x \) and \( y \) hash to the same value.

Problem 3 (CLRS 11.4-2). (2 points) Write pseudocode for hash-delete as outlined in the text, and modify hash-Insert to handle the special value Deleted. (Note that this applies to resolving collisions with open addressing.)

Problem 4 (CLRS 11.2). (5 points) Suppose that we have a hash table with \( n \) slots, with collisions resolved by chaining, and suppose that \( n \) keys are inserted into the table. Each key is equally likely to be hashed to each slot. Let \( M \) be the maximum number of keys in any slot after all the keys have been inserted. Your mission is to prove an \( O(\lg n / \lg \lg n) \) upper bound on \( E[M] \), the expected value of \( M \).

(a) Argue that the probability \( Q_k \) that exactly \( k \) keys hash to a particular slot is given by

\[
Q_k = \left( \frac{1}{n} \right)^k \left( 1 - \frac{1}{n} \right)^{n-k} \binom{n}{k}.
\]

(b) Let \( P_k \) be the probability that \( M = k \), that is, the probability that the slot containing the most keys contains \( k \) keys. Show that \( P_k \leq nQ_k \).

(c) Use Stirling’s approximation, equation (3.18) in CLRS, to show that \( Q_k < e^k / k^k \).

(d) Show that there exists a constant \( c > 1 \) such that \( Q_{k_0} < 1/n^2 \) for \( k_0 = c \lg n / \lg \lg n \). Conclude that \( P_k < 1/n^2 \) for \( k \geq k_0 = c \lg n / \lg \lg n \).

(e) Argue that

\[
E[M] \leq \Pr \left\{ M > \frac{c \lg n}{\lg \lg n} \right\} \cdot n + \Pr \left\{ M \leq \frac{c \lg n}{\lg \lg n} \right\} \cdot \frac{c \lg n}{\lg \lg n}.
\]

Conclude that \( E[M] = O(\lg n / \lg \lg n) \).