Problem 1. Let H_n denote the n-th harmonic number. Prove that

$$H_n + H_{n-1} + \cdots + H_2 + H_1 = \Theta(n \log n).$$

Problem 2.   Given an array A of length n (chosen from some set that has an underlying ordering), you can select the largest element of the array by first setting $L = A[1]$ and then comparing L to the remaining elements of the array, one a a time, replacing L with $A[i]$ if $A[i]$ is larger than L. Assume that the elements of A are all distinct and randomly chosen. What is the expected number of times you assign a value to L?

Problem 3.   What is the probability that a randomly selected bit string of length 10 is a palindrome (i.e., the bit string remains the same after reversal)?

Problem 4.   Show that if m is a positive integer, then the probability that the m-th success occurs on the $(m+n)$-th trial when independent Bernoulli trials, each with probability p of success and q of failure, are run, is $\binom{n+m-1}{n} q^n p^m$.

Problem 5.   Consider a sequence of n independent Bernoulli trials, where p denotes the probability of success in any one trial. A *run* is a maximal sequence of successes in such a sequence. For example, in the sequence $S, S, S, F, S, S, F, F, S$, where S represents success and F represents failure, there are three runs consisting of three successes, two successes, and one success, respectively. Let R denote the random variable that counts the number of runs in this sequence. Find $E(R)$ as a function of n and p.

Problem 6.   Provide an example that shows that the variance of the sum of two random variables is not necessarily equal to the sum of their variances when the random variables are not independent.

Problem 7.   Suppose that c children are choosing from among ample supplies of d different kinds of candy, with one package for each child and all choices equally likely.

(a) What is the probability that a given variety of candy is chosen by no child?

(b) What is the expected number of kinds of candy chosen by no child?

(c) Suppose that $c = d$. What happens to the expected number of kinds of candy chosen by no child?

Problem 8. Suppose that n balls are tossed into b bins so that each ball is equally likely to fall into any of the bins and that the tosses are independent.

(a) Find the probability that a particular ball lands in a specified bin.

(b) What is the expected number of balls that land in a particular bin?

(c) What is the expected number of balls tossed until a particular bin contains a ball?

(d) Consider a modified experiment in which balls are tossed until all bins contain a ball. What is the expected number of balls tossed in this experiment?

Problem 9. Let X be a random variable with mean μ and standard deviation σ. Show that for any positive real α,

$$p(|X - \mu| \geq \alpha\sigma) \leq 1/\alpha^2.$$

This inequality is called *Chebyshev's inequality*. (*Hint:* Use Markov's inequality, given in Tutorial Problem 4.)

Problem 10. Recall the hatcheck problem studied in class. Use Chebyshev's inequality to show that the probability that more than 20 people get the correct hat back when a hatcheck person returns hats at random does not exceed $1/400$ no matter how many people check their hats.

Problem 11. Suppose that the number of cans of soda pop filled in a day at a bottling plant is a random variable with an expected value of 10,000 and a variance of 1000.

(a) Use Markov's inequality to obtain an upper bound on the probability that the plant will fill more than 11,000 cans on a particular day.

(b) Use Chebyshev's inequality to obtain a lower bound on the probability that the plant will fill between 9000 and 11,000 cans on a particular day.

Problem 12. Suppose that a Bayesian spam filter is trained on a set of 500 spam messages and 200 messages that are not spam. The word "exciting" appears in 40 spam messages and in 25 messages that are not spam. Would an incoming message be rejected as spam if it contains the word "exciting" and the threshold for rejecting spam is 0.9?