1. Suppose you have access to a source that generates independent random bits. Each bit generated by the source equals 1 with probability p.

   Consider the following game: In each round, you generate two bits using your random source and compute the AND of those two bits. If the AND is 1, the game immediately ends; otherwise, you continue to the next round. Let X denote the number of ***rounds*** until the game ends.

   (a) What is the exact value of $E[X]$?

   (b) What is the exact value of $E[X^2]$? *[Hint: Use a recursive definition of X.]*

   Prove that both your answers are correct.

2. Let $U = \{1, 2, \ldots, n\}$ be a universe of n elements. Let $S_1, \ldots, S_m$ be m subsets of U, each with size $|S_i| = n/10$. We call another set of elements $X \subseteq U$ a *covering set* if X contains at least one element of each subset S_i.

   Suppose we randomly generate X by independently including each element of U with probability $p = (c \log m)/n$, for some constant $c \geq 1000$.

   (a) ***Prove*** that for any fixed index $i \in \{1, \ldots, m\}$, the set X intersects S_i with probability at least $1 - 1/m^3$. *[Hint: First compute the probability exactly, then use the world's most useful inequality.]*

   (b) ***Prove*** that X contains at most $c^2 \log m$ elements with probability at least $1 - 1/m^2$.

   (c) ***Prove*** that X is a covering set of size $O(\log m)$ with probability at least $1 - 1/m$.

   You may choose any constant value $c \geq 1000$ to make your calculations easier.

3. You are planning an election in a city with n voters and m polling stations. To minimize waiting times at the polling stations, you need to assign each voter to a single polling station where they can cast their vote. To comply with local voting laws, each voter must be ~~allowed to vote at~~ assigned a polling station within 5 miles of their residence, but at most 1000 voters can vote at any single polling station.

   You have access to a function CloseEnough(i, j) that returns True if voter i lives within 5 miles of polling station j, and returns False otherwise, in constant time. Describe an algorithm that either finds a legal assignment of polling stations to voters, or correctly reports that no such assignment exists.

Question 4 is on the other side of this page.

4. Recall that a *4CNF formula* is a boolean formula in conjunctive normal form with four literals per clause. Specifically:

   - A *literal* is either a variable x_i or its negation $\neg x_i$.
   - A *clause* is the disjunction (OR) of exactly four literals using four distinct variables.
   - Finally, a 4CNF formula is the conjunction (AND) of one or more clauses.

   For example, the following expression is a 4CNF formula with five variables and four clauses:

   $$(\neg x_1 \vee x_2 \vee x_3 \vee x_5) \wedge (x_1 \vee \neg x_2 \vee x_4 \vee x_5) \wedge (\neg x_1 \vee \neg x_3 \vee x_4 \vee x_5) \wedge (x_2 \vee x_3 \vee \neg x_4 \vee x_5).$$

   This formula evaluates to TRUE under the assignment

   $$x_1 = \text{FALSE},\ x_2 = \text{TRUE},\ x_3 = \text{TRUE},\ x_4 = \text{TRUE},\ x_5 = \text{FALSE}.$$

   Suppose you are given a 4CNF formula with n variables $x_1, \dots, x_n$ and $m \geq 100$ clauses.

   (a) Suppose we independently assign each variable x_i to be TRUE or FALSE with equal probability. A clause is *satisfied* by this assignment if the clause evaluates to TRUE; otherwise, the clause is *unsatisfied*. What is the exact expected number of ***unsatisfied*** clauses under this assignment?

   [Hint: Express the number of unsatisfied clauses in terms of indicator variables.]

   (b) ***Prove*** that the probability that at least $m/16 + 1$ clauses are ***unsatisfied*** is at most $1 - C/m$ for some constant C. *[Hint: You may use that $\frac{1}{1+t} \geq \frac{1}{2t}$ for all $t \geq 1$.]*

   (c) Part (b) implies that under a random assignment, the number of ***satisfied*** clauses is at least $15m/16$ with probability at least C/m. Using this fact, describe an *efficient* randomized algorithm that *always* finds an assignment that ***satisfies*** at least $15m/16$ clauses, and analyze its expected running time.