1. Recall the (somewhat strange) person from the first midterm who repeatedly flips a fair coin, taking a step forward when it lands head up and taking a step back when it lands tail up. Suppose this person flips the coin n times. Let $-n \leq X \leq n$ be a random variable indicating the position of the walker after n steps.

   a. [5 points] What is $\mathsf{E}[X]$?
   Let H be the number of times the coin lands head up and let T be the number of times it lands tail up. Then $X = H - T$ so $\mathsf{E}[X] = \mathsf{E}[H - T] = \mathsf{E}[H] - \mathsf{E}[T] = n/2 - n/2 = 0$. $\mathsf{E}[H] = n/2$ since H is a binomial$(n, 1/2)$ random variable. So is T.

   b. [5 points] What is $\mathsf{Var}[X]$?
   Notice that since $H + T = n$, $X = H - (n - H) = 2H - n$. Thus $\mathsf{Var}[X] = \mathsf{Var}[2H - n] = 4\mathsf{Var}[H] = n$, again using the fact that H is a binomial$(n, 1/2)$ random variable.

   c. [5 points] What is the probability distribution of X?
   $$P(X = x) = P(2H - n = x) = P\left(H = \frac{x+n}{2}\right) = \begin{cases} \binom{n}{(x+n)/2} 2^{-n} & \text{if } x + n \text{ even;} \\ 0 & \text{otherwise.} \end{cases}$$

   d. [10 points] For $n = 10,000$, estimate the probability that the walker ends up more than 100 steps from the starting position.
   From part (b), $\mathsf{SD}[X] = 100$. Since n is large and the probabilities in (c) are given by a binomial distribution, they are well approximated by a normal distribution with mean 0 and standard deviation 100. The probability of being within 1 standard deviation of the mean is approximately 68% for a normal distribution, so the probability of being more than 100 steps from the starting position is approximately 32%.

	score
1	
2	
3	
4	
EC	
total	

2. The first three picks in the NBA draft are determined by a lottery: Fourteen balls labeled 1 to 14 are placed in an urn and four balls are drawn, without replacement. 250 of the possible combinations are assigned to the team with the worst season record, 199 to the team with the second worst season record, *etc.* One combination, $(11, 12, 13, 14)$, is not assigned to any team; if it is drawn the balls are returned to the urn and the process is repeated. The team to which the combination of numbers drawn is assigned gets the first pick in the draft.

   a. [5 points] How many different combinations of numbers can be drawn?
   $\binom{14}{4} = \frac{14 \cdot 13 \cdot 12 \cdot 11}{1 \cdot 2 \cdot 3 \cdot 4} = 7 \cdot 13 \cdot 11 = 1001.$

   b. [10 points] What is the probability that the team with the second worst season record gets the first pick in the draft?
   For the second worst team to get the first pick, one of its 199 combinations out of the $1001 - 1 = 1000$ combinations that are assigned to teams. So the probability is 0.199.

   c. [10 points] After the first pick is determined, the balls are returned to the urn, and the process is repeated. If a new team (*i.e.*, not the one that won the first pick) wins, it gets second pick; otherwise the balls are returned to the urn and the process is repeated. What is the probability that the team with the worst season record gets the second pick in the draft?
   For this team to get the second pick, one of the other teams must win the first pick, and then this team must win the second pick. Let the number of combinations assigned to the team with the i^{th} worst record be c_i, and let n be the number of teams in the lottery. Since the events that the team with the i^{th} worst record wins the first pick are disjoint, the probabilities add, so the probability that the worst team gets the second pick is:

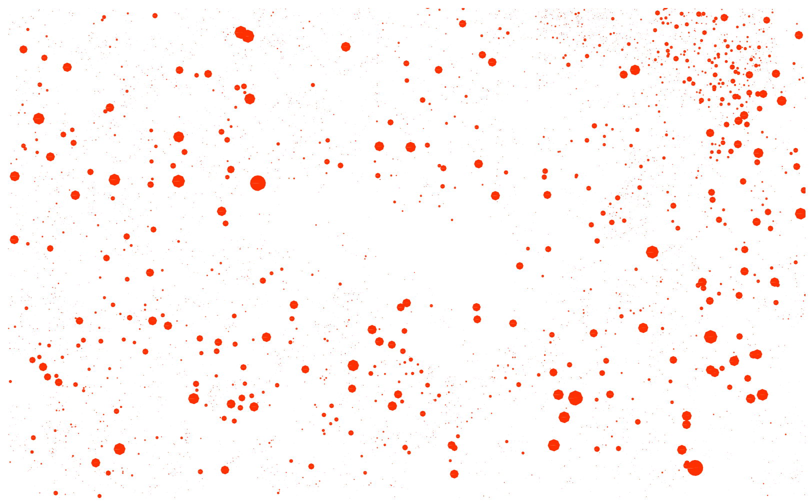
$$\sum_{i=2}^{n} \frac{c_i}{1000} \cdot \frac{250}{1000 - c_i}.$$

3. A gambler offers to pay you 2^n dollars if you flip a fair coin until it lands head up, and it does so on the n^{th} flip.

   a. [15 points] What is the number of dollars you expect to be paid if you play this game?
   Let $W = 2^N$ be the number of dollars the gambler pays you, where N is the number of times you flip the coin until it lands head up. Then $\mathsf{E}[W] = \sum_{n=1}^{\infty} 2^n P(N = n) = \sum_{n=1}^{\infty} 2^n 2^{-n} = \sum_{n=1}^{\infty} 1 = \infty$.

   b. [10 points] How much would you pay to play this game? Explain your answer.
   I accepted several different answers for this problem. The most conservative strategy is to pay no more than \$2 to play, since then you are assured of winning some money, even if the first head appears on the first flip. Less risk-averse players could pay any finite amount of money to play this game and still have a positive expected net payoff.
   The traditional analysis of this "St. Petersburg paradox" introduces the notion of *utility*, which is an increasing function $u : \mathbb{R} \to \mathbb{R}$, but with $u'' < 0$, *e.g.*, $u(w) = \log w$. This is supposed to capture the idea that if you have only \$1000, another \$1000 has larger utility than if you start with \$1,000,000. The argument then is that it is $\mathsf{E}[u(W)]$ that is relevant, not $\mathsf{E}[W]$. Using the logarithm makes the infinite sum defining the expectation value of the utility finite, in this problem. This is not completely satisfactory, since for any increasing utility function, there is some set of winnings that make the expectation value infinite.

4. There is a $500m. \times 1000m.$ plot of land on Barro Colorado Island in Panama where a careful census has been made of the trees. The locations of trees of one common species, *Alseis blackiana*, are indicated by dots in the graphic below, where the size of each dot represents the size of the tree.



a. [10 points] Do you think the locations of these trees are a Poisson scatter? Why or why not?
Looking at the graphic, there appear to be regions with greater and lesser densities of trees, which suggests that their locations are not a Poisson scatter. I also accepted the observation that since a tree cannot be located within the trunk of another tree, there is not equal probability of a point being located in every region of the same area, so the locations do not satisfy the defining condition of a Poisson scatter. Or you could postpone answering this part until after part (b).

b. [15 points] There are 7599 dots in this graphic. If we partition the plot into four congruent rectangles, by dividing it in half vertically and horizontally, the number of dots in each rectangle is 2115, 1950, 1708 and 1826, moving counterclockwise from the northeast rectangle. What is the probability of observing this distribution if the tree locations are a Poisson scatter?
If this were a Poisson scatter, the intensity per area of size $250m. \times 500m.$ would be approximately $7599/4 \approx 1900$, and the number of dots per rectangle would be i.i.d. Poisson random variables, so the probability of observing these numbers of trees would be approximately:

$$\frac{1900^{2115}}{2115!}e^{-1900} \cdot \frac{1900^{1950}}{1950!}e^{-1900} \cdot \frac{1900^{1708}}{1708!}e^{-1900} \cdot \frac{1900^{1826}}{1826!}e^{-1900}.$$

This is a really, really, really small number, since the numbers of dots in the rectangles are not close to 1900, measured in units of the standard deviation, $\sqrt{1900} \approx 44$.

5. [Extra credit] In this problem "simulate" means "do an experiment with outcomes having the same probabilities as the outcomes of".

   a. [5 points] Suppose you have a coin with unknown bias. How can you simulate the flip of a fair coin?
   Flip the coin twice. If the sequence of outcomes is head-tail, call it "Head"; if it is tail-head, call it "Tail". If it is neither of these, repeat. Since the probability of head-tail is pq is the probability of tail-head, the events we are calling "Head" and "Tail" occur equally often, so this simulates a fair coin flip.

   b. [10 points] Suppose you have a fair coin. How can you simulate the flip of a biased coin that lands head up with probability 1/3? [Hint: Think about conditional probabilities.]
   Flip the coin three times. If it never lands head up or it lands head up more than once, repeat. If the head occurs on the first of the three flips, call that "Head"; call it "Tail" if the head occurs on the second or third flip. Since each has the same probability, the probability of "Head" is 1/3.