1. Suppose there are N total tuna in the seas surrounding Alaska. You gently capture, tag and release k of them. A week later you come back and catch r tuna. There are m of these tuna with one of your tags (and $r - m$ without).

   (a) What is the probability, p_N, of this event?

   Solution: $p_N = \frac{C_{k,m} C_{N-k,r-m}}{C_{N,r}}$.

   (b) If we were to solve $p_N / p_{N-1} \geq 1$ (see Example 2.42) we would find this is true only when $N \leq kr/m$. Say in words what this means.

   Solution: Whenever the ratio p_N / p_{N-1} is larger than 1 we must have p_N is increasing. This stops occurring when $N > kr/m$. Thus, $N = kr/m$ is the maximum value of p_n.

   (c) Notice that $k/(kr/m) = m/r$. Explain why having these two ratios equal makes sense as a guess for N.

   Solution: Both ratios have the same proportion of tuna present.

   (d) In order to estimate the size of the population of bluegills (a species of fresh water fish) in a small lake in Missouri, a total of 250 bluegills are captured and tagged and then released. After allowing sufficient time for the tagged fish to disperse, a sample of 150 bluegills were caught. It was found that 16 bluegills in the sample were tagged. Estimate the size of the bluegill population in this lake.

   Solution: $k = 250, r = 150, m = 16$ so $N = 250 \cdot 150/16 = 2343$ is our best guess.

2. This question is about a concept called *thinning*—where we take a random variable and make it smaller.

   (a) Fix numbers $0 < q, p < 1$ and a positive integer $n > 0$. Let $Y = \text{binomial}(n, p) = \sum_1^n X_i$ with $X_i = \text{bernouli}(p)$. Suppose $W = \text{binomial}(Y, q)$ and $X_i' = \text{bernouli}(q)$ for $i = 1, 2, \ldots, n$. Use the random variables X_i and X_i' to explain why $W = \text{binomial}(n, pq)$. *Hint: Set $\tilde{X}_i = X_i X_i'$. What is the distribution of $\tilde{X}_i$?*

   Solution: We can write $W = \sum_{i=1}^n X_i X_i'$. If we call $\tilde{X}_i = X_i X_i'$ notice that $P(\tilde{X}_i = 1) = pq$, and the only other option is $P(\tilde{X}_i = 0) = 1 - pq$. Thus $\tilde{X}_i = \text{bernouli}(pq)$. It follows that

   $$W = \sum_{i=1}^n \tilde{X}_i = \text{binomial}(n, pq).$$

   (b) You are picking blackberries on a sunny day. It is bright, so you can't see the berries that well. You pick 100 of them. Each is ripe independently with probability 9/10. Each has a bit of spiderweb on it with probability 1/9. The presence of a bit of web is independent of the ripeness and webs on other berries. What is the probability you pick at least 98 ripe berries with no spiderwebs?

   Solution: By the previous part the number of ripe, webless berries is

   $$N = \text{binomial}(\text{binomial}(100, 9/10), 8/9) = \text{binomial}(100, 4/5).$$

   We then have

   $$P(N \geq 98) = C_{100,98}.8^{98}.2^2 + C_{100,99}.8^{99}.2 + C_{100,98}.8^{100}.$$

   (c) Set $Z = \text{Poisson}(\lambda)$ and suppose $M = \text{binomial}(Z, q)$. We can't prove it with the material covered in this class, but an important property called *Poisson thinning* is that $M = \text{Poisson}(\lambda q)$. The United States have on average 20 shark attacks per year. The probability of a shark attack being fatal is 1/5. What is the probability of at least 3 deaths by shark attack this year in the United States?

   Solution: The number of deaths is distributed like $D = \text{Poisson}(20 * 1/5) = \text{Poisson}(4)$. We have

   $$P(D \geq 3) = 1 - P(D < 3) = 1 - e^{-4} - 4e^{-4} - 16e^{-4}/2 = .76.$$

   (d) The most amazing thing (in my opinion) about the Poisson random variable is that if we start with $Z = \text{Poisson}(\lambda)$, then apply a multinomial thinning of Z with parameters $p_1, \ldots, p_n$ then the number put in group i is distributed like $Z_i = \text{Poisson}(\lambda p_1)$. Moreover, the Z_i are **independent**. No other random variable has this independence property! Suppose we place $\text{Poisson}(n)$ balls uniformly randomly into n bins. What is the distribution for the number of balls inside of the first bin?

   Solution: $Z_1 = \text{Poisson}(n(1/n)) = \text{Poisson}(1)$

   (e) Let Q be the number of empty bins after placing all $\text{Poisson}(n)$ balls. Explain why Poisson thinning ensures $Q = \text{binomial}(n, p)$, and find p.

   Solution: The number of balls in each bin is $\text{Poisson}(1)$ and is independent by Poisson thinning. The probability a bin is empty is $P(Z_1 = 0) = e^{-1}$. Hence, $p = 1/e$.

   (f) Suppose n balls are placed uniformly into n bins. Find the probability, q, that the first bin is empty after this. Explain why the number of empty bins is not a $\text{binomial}(n, q)$ random variable.

> **Solution:** $q = (1 - 1/n)^n$. The number of balls in a given bin is dependent on the others.

3. You keep playing a game with a 9/10 chance of losing \$1 and a 1/10 chance of winning a dollar. Let $X_i \in \{-1, 1\}$ be the amount of money you win on the ith try and $S_n = \sum_{i=1}^{n} X_i$ be your total winnings/losses.

   (a) Let E_n be the event that $S_{2n} = 0$. What is $P(E_n)$?

   > **Solution:** This can only happen if the the number of wins is equal to the number of losses. So we have
   > $$P(E_n) = C_{2n,n}(1/10)^n(9/10)^n = C_{2n,n}(9/100)^n.$$

   (b) Let A be the event that there exists a time t such that $S_t = 0$. Write A in terms of the events E_n, and explain why.

   > **Solution:** $A = \cup_{n=1}^{\infty} E_n$. This is because the only way S_t can be 0 is if at least one of the events E_n occur. Note $S_t \neq 0$ if t is odd.

(c) Use a union bound and the previous part to give an upper bound on $P(A)$.

Solution: $P(A) \leq \sum_{n=1}^{\infty} C_{2n,n}(9/100)^n$.

(d) Use the fact that $C_{2n,n} \leq 4^n$ and that $\sum_{n=1}^{\infty} a^n = \frac{1}{1-a} - 1$ to estimate the sum in the previous part.

Solution:

$$P(A) \leq \sum_{n=1}^{\infty} 4^n (9/100)^n = \sum_{n=1}^{\infty} (36/100)^n = 36/64 \approx .56.$$

(e) Say in words what it means that $P(A) < 1$. Is this surprising to you?

Solution: This says that there is a positive probability that you never get back to \$0. Even if you play the game forever. Many find it surprising that despite infinite tries there is a chance of never making your money back.

Answers to Homework 2

4. $P_{30,5}$
9. (a) 17576000 (b) .63
16. .1
21. (a) 120 (b) 60 (c) 50400 (d) 34650
24. .054
34. .3548
39. .3297
47. .4232
56. 14/33
61. (a) .41 (b) .51 (c) .077
68. (a) .07 (b) Two seniors, two juniors and one sophomore
81. .3439