[Homework] Expectation

1. Let X be the number of heads minus the number of tails obtained when a fair coin is tossed n times. Calculate the mean of X.

   Sol. Let i be the number of tails.

   Then $n - i$ is the number of heads and $X = n - 2i$.

   $$P(X = i) = \binom{n}{i}(\frac{1}{2})^i(\frac{1}{2})^{n-i} = \binom{n}{i}(\frac{1}{2})^n$$

   $$\begin{aligned} E(X) &= \sum_x x \cdot P(X = x) \\ &= \sum_{i=0}^{n}(n - 2i)\binom{n}{i}(\frac{1}{2})^n \\ &= \{n\sum_{i=0}^{n}\binom{n}{i} - 2\sum_{i=0}^{n}i \cdot \binom{n}{i}\} \cdot (\frac{1}{2})^n \\ &= \{n \cdot 2^n - 2 \cdot n \cdot 2^{n-1}\} \cdot (\frac{1}{2})^n \\ &= 0 \cdot (\frac{1}{2})^n = 0 \end{aligned}$$

2. A salesman has scheduled two appointments to sell encyclopedias. His first appointment will lead to a sale with probability 0.3, and his second will lead independently to a sale with probability 0.6. Any sale made is equally likely to be either for the deluxe model, which costs \$1000, or the standard model, which costs \$500. Determine the expectation of X, the total dollar value of all sales.

   Sol. In Homework4 #4, The possible values of X are:

   2000 dollars. In this case, we have

   $$P\{X = 2000\} = P(S_1 \cap D_1)P(S_2 \cap D_2) = \frac{0.3}{2} \cdot \frac{0.6}{2} = 0.045$$

   1500 dollars. In this case, we have

   $$P\{X = 1500\} = P(S_1 \cap D_1)P(S_2 \cap D_2^c) + P(S_1 \cap D_1^c)P(S_2 \cap D_2) = 0.09$$

   1000 dollars. In this case, we have

   $$P\{X = 1000\} = P(S_1 \cap D_1)P(S_2^c) + P(S_1^c)P(S_2 \cap D_2) + P(S_1 \cap D_1^c)P(S_2 \cap D_2^c) = 0.315$$

500 dollars. In this case, we have

$$P\{X = 500\} = P(S_1 \cap D_1^c)P(S_2^c) + P(S_1^c)P(S_2 \cap D_2^c) = 0.27$$

0 dollars. In this case, we have

$$P\{X = 0\} = P(S_1^c)P(S_2^c) = 0.28$$

$$EX = (2000 \cdot 0.045) + (1500 \cdot 0.09) + (1000 \cdot 0.315) + (500 \cdot 0.27) + (0 \cdot 0.28) = 675 \text{dollars}.$$

3. One of the numbers from 1 through 10 is randomly chosen. You are to try to guess the number chosen by asking "yes-no" questions. Compute the expected number of questions you will need to ask in each of the following two cases:

(a) Your i^{th} question is "Is it i?" $i = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$.

Sol. If at stage n by asking the question "is it i?", one is able to eliminate one possible choice from further consideration (assuming that we have not guessed the correct number)before stage n. Thus let E_n be the event at stage n we guess the number correctly, assuming we have not guessed it correctly in the $n - 1$ earlier stages. Then

$$P(E_n) = \frac{1}{10 - (n - 1)} = \frac{1}{11 - n}$$

so we have that

$$P(E_1) = \frac{1}{10}, P(E_2) = \frac{1}{9}, P(E_3) = \frac{1}{8}, \cdots, P(E_{10}) = 1$$

The expected number of guesses to make using this method is then given by

$$
\begin{aligned}
E(N) &= 1 \cdot \frac{1}{10} + 2 \cdot (1 - \frac{1}{10}) \cdot \frac{1}{9} + 3 \cdot (1 - \frac{1}{10}) \cdot (1 - \frac{1}{9}) \cdot \frac{1}{8} + \cdots \\
&= 1 \cdot \frac{1}{10} + 2 \cdot \frac{1}{10} + \cdots + 10 \cdot \frac{1}{10} \\
&= 5.5
\end{aligned}
$$

(b) With each question you try to eliminate one-half of the remaining numbers, as nearly as possible.

Sol. Divide what you have in halves and ask, "is it greater than or equal to x?" where x is the middle of the numbers. For instance, your first question is, "is it greater than or equal to 5.5?". Let X denote the random number, so that $P\{X = i\} = \frac{1}{10}$ for $i = 1, \cdots, 10$.

If $X = 1, 2, 3, 6, 7, 8$ then Q = 3.

If $X = 4, 5, 9, 10$ then Q = 4.

Therefore, $P\{Q = 3\} = P\{X = 1\} + P\{X = 2\} + P\{X = 3\} + P\{X = 6\} + P\{X = 7\} + P\{X = 8\} = 0.6$, and $P\{Q = 4\} = 0.4$.

Thus, $E(Q) = (3 \cdot 0.6) + (4 \cdot 0.4) = 3.4$.

4. Two baseball teams A and B with equal skills play korean series, where the team who wins 4 games first becomes the champion. What is the expected number of games to play?

Sol. Let X be the number of games, then $X \sim NB(4, \frac{1}{2})$.

$$P(X = 4) = 2 \cdot \binom{4-1}{4-1}(1 - 0.5)^{4-4}(0.5)^4 = \frac{1}{8}$$

$$P(X = 5) = 2 \cdot \binom{5-1}{4-1}(1 - 0.5)^{5-4}(0.5)^4 = \frac{1}{4}$$

$$P(X = 6) = 2 \cdot \binom{6-1}{4-1}(1 - 0.5)^{6-4}(0.5)^4 = \frac{5}{16}$$

$$P(X = 7) = 2 \cdot \binom{7-1}{4-1}(1 - 0.5)^{7-4}(0.5)^4 = \frac{5}{16}$$

$$E(X) = 4 \cdot \frac{1}{8} + 5 \cdot \frac{1}{4} + 6 \cdot \frac{5}{16} + 7 \cdot \frac{5}{16} = \frac{93}{16}$$

5. A player throws a fair die and simultaneously flips a fair coin. If the coin lands heads, she wins \$2. If tails, she wins one-half of the value that appears on the die. Determine her expected winnings.

Sol. Let $H = \{\text{heads}\}$, and define N to be the number of dots on the rolled die. We know that N and H are independent. Let W denote the amount won. We know that

$$P(W = 2|H) = 1 \text{and} P(W = N/2|H^c) = 1.$$

Therefore,

$$
\begin{aligned}
E(W) &= E(W|H)P(H) + E(W|H^c)P(H^c) \\
&= E(2|H)P(H) + E(\frac{N}{2}|H^c)P(H^c) \\
&= 2P(H) + E(\frac{N}{2})P(H^c)
\end{aligned}
$$

by independence. But $P(H) = P(H^c) = \frac{1}{2}$, and $E(N) = \frac{1+\cdots+6}{6} = \frac{7}{2}$.

Consequently,

$$E(W) = 2P(H) + \frac{1}{2}E(N)P(H^c) = \frac{15}{8}$$

6. Two gamblers with equal skills play games. The player will win or lose \$1 for each game. Let W be the net winnings of a gambler whose strategy is to stop gambling immediately after his first win. Find

   (a) $P(W > 0)$.
      Sol. There is only one way that will happen is when he wins the first time. Hence, $P(W > 0) = P(\text{win}) = \frac{1}{2}$

   (b) $P(W < 0)$.
      Sol. Since $P(W = 0) = P(\text{lose,win}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ and

   $$P(W > 0) + P(W = 0) + P(W < 0) = 1$$

   So we have
   $$P(W < 0) = \frac{1}{4}$$

   (c) EW.
      Sol. Let X be the number of trial of which winning occurs.

   $$W = 1 \cdot 1 + (-1) \cdot (X - 1)$$

   $$
   \begin{aligned}
   E(W) &= E(E(W|X)) \\
   &= \sum_x E(W|X = x)P(X = x) \\
   &= \sum_{x=1}^{\infty}(2 - x) \cdot \frac{1}{2^x} \\
   &= 0
   \end{aligned}
   $$

7. A box contains five tickets labeled with numbers $-3, -1, 0, 1, 3$. We do 100 random draws with replacement from the box. Calculate the expected value of the sum of the positive numbers on the tickets drawn.
   Sol. Each of the five numbers should be drawn about $\frac{100}{5} = 20$ times. Taking just the positive numbers, 1 and 3,so the expected value of the sum of the positive numbers on the tickets drawn is

   $$1 \cdot 20 + 3 \cdot 20 = 80$$

8. There are n components lined up in a linear arrangement. Suppose that each component independently functions with probability p. Let X be the number of

2 neighboring pairs that both are functional. Calculate EX.

Sol. Let a_i be the ith component and X_i be the a_i and a_{i+1} are functional.

$$X_i = \begin{cases} 1 & a_i \text{ and } a_{i+1} \text{ are functional with probability } p^2 \\ 0 & \text{otherwise with probability } 1 - p^2 \end{cases}$$

Then,

$$
\begin{aligned}
EX &= E(\sum_{i=1}^{n-1} X_i) \\
&= \sum_{i=1}^{n-1} EX_i \\
&= \sum_{i=1}^{n-1} \{1 \cdot p^2 + 0 \cdot (1 - p^2)\} \\
&= (n-1)p^2
\end{aligned}
$$

9. A box contains 5 red and 5 blue marbles. Five marbles are withdrawn randomly with replacement. Each time you have red you win \$1.10 and each time you have blue you lose \$1.00. Let X be the total amount you win after five trials.

   (a) Let R be the number of red marbles drawn during 5 draws. Represent X as a function of R.
   Sol.
   $$X = 1.1R - 1(5 - R) = 2.1R - 5$$

   (b) Calculate EX.
   Sol. Since $R \sim B(5, \frac{1}{2})$, $ER = np = \frac{5}{2}$
   $$EX = E(2.1R - 5) = 2.1ER - 5 = \frac{1}{4}$$

10. Four buses carrying 148 students from the same school arrive at a football stadium. The buses carry, respectively, 40, 33, 25, and 50 students. One of the students is randomly selected. Let X denote the number of students that were on the bus carrying the randomly selected student. One of the 4 bus drivers is also randomly selected. Let Y denote the number of students on her bus.

   (a) Which of EX and EY do you think is larger? Why?
   Sol.

(b) Compute EX and EY.

Sol. X corresponds to the method of drawing a student at random. The probability that a student comes from bus i, $i = \{A, B, C, D\}$ is simply the number of students on that bus divided by the total number of students.

x	40	33	25	50	
$P(X = x)$	$\frac{40}{148}$	$\frac{33}{148}$	$\frac{25}{148}$	$\frac{50}{148}$	1

Then,

$$E(X) = 40 \cdot \frac{40}{148} + 33 \cdot \frac{33}{148} + 25 \cdot \frac{25}{148} + 50 \cdot \frac{50}{148} = 39.28.$$

Y corresponds to the method of drawing a bus driver at random. The probability that a bus driver comes from a particular bus is constant, $\frac{1}{4}$.

y	40	33	25	50	
$P(Y = y)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

Then,

$$E(Y) = 40 \cdot \frac{1}{4} + 33 \cdot \frac{1}{4} + 25 \cdot \frac{1}{4} + 50 \cdot \frac{1}{4} = 37$$

.

11. There are n components lined up in a linear arrangement. Suppose that each component independently functions with probability p. Let X be the number of 2 neighboring pairs that both are not functional. Calculate $P(X = 0)$.

Sol 1. Let $X=$ the number of nonfunctional components and $E=$event that no two neighboring components are both functional . Given n components with m functional, there are $\binom{n-m+1}{m}$ possible ordering in which no two neighboring components are both nonfunctional provided $n + 1 \geq 2m$.

If $n+1 < 2m$, then there must be a pair of neighboring nonfunctional components.

$$
\begin{aligned}
P(E) &= \sum_{m=0}^{n} P(E|X = m)P(X = m) \\
&= \sum_{0 \leq m \leq \frac{n+1}{2}} \frac{\binom{n-m+1}{m}}{\binom{n}{m}} \cdot \binom{n}{m} p^{n-m}(1-p)^m \\
&= \sum_{0 \leq m \leq \frac{n+1}{2}} \binom{n-m+1}{m} \cdot \binom{n}{m} p^{n-m}(1-p)^m
\end{aligned}
$$

Sol 2.
$$P_n = p \cdot P_{n-1} + p \cdot (1-p) \cdot P_{n-2}, \ P_1 = 1, P_2 = 2p - p^2$$

$$x^2 - px + p(p-1) = 0, \alpha, \beta = \frac{p \pm \sqrt{4p - 3p^2}}{2}$$

$$P_n = a\alpha^n + b\beta^n$$

$$P_n = \frac{2p^2 - 3p - \sqrt{4p - 3p^2}}{-p\sqrt{4p - 3p^2} - (4p - 3p^2)}\alpha^n + \frac{-2p^2 + 3p - \sqrt{4p - 3p^2}}{-p\sqrt{4p - 3p^2} + 4p - 3p^2}\beta^n$$

Extra

1. A statistics class contains 368 students. Homework in the class is submitted online and graded automatically. The instructor wrote software to detect cheating on the homework. Suppose that the software has a 99.6% chance of correctly identifying a student who cheats, and a 0.2% chance of mis-identifying an honest student as a cheater. Assume that the software identification of students as cheaters or honest is independent from student to student, and that twelve of the students cheat on the homework.

   (a) A student is selected at random from the class. Compute the probability that the student cheated, given that the software says he or she did.

   (b) Compute the expected number of cheaters who are correctly identified by the software.

   (c) Compute the expected number of honest students who are incorrectly identified to be cheaters by the software.

   (d) Compute the probability that the software correctly identifies at least ten of the cheaters.

   (e) Compute the probability that the software correctly identifies all twelve of the students who cheated, without mis-identifying any of the 356 honest students to be cheaters.